

34437

National Library  
of CanadaBibliothèque nationale  
du CanadaCANADIAN THESES  
ON MICROFICHETHÈSES CANADIENNES  
SUR MICROFICHE

NAME OF AUTHOR/NOM DE L'AUTEUR

ROBERT E. MERCER

TITLE OF THESIS/TITRE DE LA THÈSE

ADAPTIVE SEARCH USING A  
REPRODUCTIVE META-PLAN

UNIVERSITY/UNIVERSITÉ

UNIVERSITY OF ALBERTA

DEGREE FOR WHICH THESIS WAS PRESENTED/  
GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE

MSc

YEAR THIS DEGREE CONFERRED/ANNÉE D'OBTENTION DE CE GRADE

1977

NAME OF SUPERVISOR/NOM DU DIRECTEUR DE THÈSE

J. R. SAMPSON

Permission is hereby granted to the NATIONAL LIBRARY OF  
CANADA to microfilm this thesis and to lend or sell copies  
of the film.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈ-  
QUE NATIONALE DU CANADA de microfilmer cette thèse et  
de prêter ou de vendre des exemplaires du film.

The author reserves other publication rights, and neither the  
thesis nor extensive extracts from it may be printed or other-  
wise reproduced without the author's written permission.

L'auteur se réserve les autres droits de publication; ni la  
thèse ni de longs extraits de celle-ci ne doivent être imprimés  
ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/DATE

15 July 1977

SIGNED/SIGNÉ

PERMANENT ADDRESS/RÉSIDENCE FIXE

383 DOWWOOD COURT

SHERWOOD PARK, ALBERTA

T8A 1K5



National Library of Canada

Cataloguing Branch  
Canadian Theses Division

Ottawa, Canada  
K1A 0N4

Bibliothèque nationale du Canada

Direction du catalogage  
Division des thèses canadiennes

## NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

**THIS DISSERTATION  
HAS BEEN MICROFILMED  
EXACTLY AS RECEIVED**

## AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

**LA THÈSE A ÉTÉ  
MICROFILMÉE TELLE QUE  
NOUS L'AVONS REÇUE**

THE UNIVERSITY OF ALBERTA

ADAPTIVE SEARCH USING A REPRODUCTIVE META-PLAN

by



Robert E. Mercer

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE  
OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTING SCIENCE

EDMONTON, ALBERTA

FALL, 1977

THE UNIVERSITY OF ALBERTA  
FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and  
recommend to the Faculty of Graduate Studies and Research,  
for acceptance, a thesis entitled ADAPTIVE SEARCH USING A  
REPRODUCTIVE META-PLAN submitted by Robert E. Mercer in  
partial fulfilment of the requirements for the degree of  
Master of Science

.....*M. Simpson*.....  
Supervisor

.....*Kellogg Wilson*.....

.....*G. K. Schubert*.....

.....*Kenneth Morgan*.....

Date.....*15 July 1977*.....

## ABSTRACT

A reproductive plan is a type of adaptive procedure devised by J. H. Holland which embodies many principles found in the adaptation of natural systems through evolution. The primary objective of this study is the development of a reproductive meta-plan, an adaptive procedure that controls the modification of certain parameter values in a lower level reproductive plan. This work also provides direction for further study in the area of meta-adaptation.

The second chapter introduces the subject of adaptation and the concept of a reproductive plan. Some reproductive plans are reviewed. Chapter 3 describes a non-reproductive meta-plan developed by D. J. Cavicchio and discusses some of its limitations. A reproductive meta-plan is developed in an attempt to remove some of these limitations.

Chapter 4 summarizes and discusses the results of experiments using both meta-plans. A limited improvement in the lower level reproductive plan's utility gain is found when the reproductive meta-plan is used. Further, the space of parameter value combinations is searched more vigorously by the reproductive meta-plan. The results also indicate potentially fruitful areas that could be intensively investigated. Chapter 5 concludes with a discussion of those and other areas for future research.

#### ACKNOWLEDGEMENTS

My sincerest thanks to my supervisor Jeff Sampson for his generous financial support and patience. His many readings of earlier drafts and his suggestions on writing style were most helpful. Thanks also goes to the other members of the committee Ken Morgan, Len Schubert, and Kelly Wilson for their many helpful suggestions and their careful reading of the thesis. This research was supported in part by the National Research Council of Canada, through Grant A-7643 to J.R. Sampson.

## TABLE OF CONTENTS

| Chapter                                                                            | Page |
|------------------------------------------------------------------------------------|------|
| 1 INTRODUCTION .....                                                               | 1    |
| 2 REPRODUCTIVE PLANS .....                                                         | 8    |
| 2.1 Holland's Framework .....                                                      | 9    |
| 2.2 Reproductive Plans .....                                                       | 13   |
| 2.2.1 Schemata .....                                                               | 14   |
| 2.2.2 General Scheme of Plans of Type <i>R</i> .....                               | 15   |
| 2.2.3 Generalized Genetic Operators .....                                          | 16   |
| 2.3 Optimal Allocation of Trials and the<br>Robustness of Reproductive Plans ..... | 20   |
| 2.4 Bagley's Results .....                                                         | 22   |
| 2.5 Cavicchio's Results .....                                                      | 24   |
| 2.6 Hollstien's Results .....                                                      | 29   |
| 2.7 Frantz's Results .....                                                         | 31   |
| 2.8 De Jong's Results .....                                                        | 33   |
| 3 A REPRODUCTIVE META-PLAN .....                                                   | 38   |
| 3.1 Meta-adaptation of Genetic Operator Values .....                               | 39   |
| 3.2 A Critique of Cavicchio's Meta-adaptive Plan .....                             | 43   |
| 3.3 A Proposed Reproductive Meta-plan .....                                        | 48   |
| 4 EXPERIMENTAL RESULTS .....                                                       | 58   |
| 4.1 Performance of Cavicchio's Meta-plan .....                                     | 60   |
| 4.2 Performance of the Proposed Reproductive<br>Meta-plan .....                    | 66   |
| 4.3 Results of an Experiment Involving a "Crisis"<br>Point .....                   | 77   |
| 4.4 Results of an Experiment Done on Breakpoints .....                             | 79   |

| Chapter                                | Page |
|----------------------------------------|------|
| 4.5 General Discussion .....           | 84   |
| 5 CONCLUSIONS AND FUTURE STUDIES ..... | 91   |
| REFERENCES .....                       | 97   |

## LIST OF TABLES

| Table |                                                                                                                                                         | Page |
|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 4.1   | Initial meta-population and meta-parents for the remaining meta-generations of an experiment using a 60/10(10) population                               | 66   |
| 4.2   | Number of (possible) new population members per meta-structure and a summary of the average utility of the 10 best structures in a 60/10(10) population | 68   |
| 4.3   | Meta-generation showing the selection method for the original and revised reproductive meta-plans                                                       | 73   |
| 4.4   | Summary of breakpoint biasing experiment using two 70/10(10) populations (one for biased meta-structures, and one for unbiased ones)                    | 83   |

## LIST OF FIGURES

| Figure |                                                                                                                         | Page |
|--------|-------------------------------------------------------------------------------------------------------------------------|------|
| 3.1    | Illustration of inappropriate action taken as a result of too many generations per meta-generation                      | 49   |
| 3.2    | Illustration of too little action taken as a result of too many generations per meta-generation                         | 49   |
| 3.3    | Method used by the reproductive plan for producing offspring. MS stands for meta-structure                              | 51   |
| 3.4    | Representation of a meta-structure                                                                                      | 53   |
| 4.1    | Genetic operator values of Cavicchio's meta-plan, with preselection, for 75 generations                                 | 61   |
| 4.2    | Genetic operator values of Cavicchio's meta-plan, without preselection, for 75 generations                              | 62   |
| 4.3    | Average utility of 10 best structures of Cavicchio's meta-plan with and without preselection                            | 63   |
| 4.4    | Genetic operator rates of original reproductive meta-plan for 16 generations                                            | 69   |
| 4.5    | Average utility of 10 best structures for original and revised reproductive meta-plans for 27 generations               | 70   |
| 4.6    | Genetic operator rates of revised reproductive meta-plan for 14 generations                                             | 76   |
| 4.7    | Genetic operator rates of Cavicchio's meta-plan with preselection, at a "crisis" point                                  | 80   |
| 4.8    | Comparison of average utility of Cavicchio's meta-plan, with preselection, at a "crisis" point and a non-"crisis" point | 81   |

Figure

Page

- 4.9      Summary of average utility of the  
         10 best structures for Cavicchio's meta-plan  
         with and without preselection and the  
         original and revised reproductive  
         meta-plans      86
- 4.10      First differences of the average  
         utilities summarized in Figure 4.9  
         (a) Reproductive meta-plan  
         (b) Cavicchio's meta-plan with preselection  
         (c) Cavicchio's meta-plan without preselection 88

## CHAPTER 1

### INTRODUCTION

Frequently a certain level of performance for a device interacting with an environment must be maintained. For example a heating unit (the device) must keep the temperature of the air in a room (the environment) within a few degrees (the level of performance) of a desired temperature. The plan to control this interaction could be a thermostat.

Other examples can be found in a variety of natural and artificial situations. An insect species or population acquiring resistance to a new insecticide is an example of control in evolution. The plan must selectively utilize the results of the genetic variance introduced by recombination and mutation. A program that has "learned" to play average chess is an example of control in an artificial system. The plan must adapt the chess program as more knowledge about the game becomes available to the plan. Evidently, there are a variety of problems, in a diversity of fields, which require control of device-environment interaction.

Most investigators (e.g., [2], [13]) acknowledge three broad classes of processes (device-environment interactions) and plans: deterministic, stochastic, and

adaptive. Deterministic plans employ a complete description of the device and the environment. The state transitions of the device and the environment may be expressed as equations which permit no variation in behaviour within repetitions of a given set of circumstances. The thermostat example is of this type. This device can be in two states, "on" and "off". The "equations" of control might be: 'when the state is "on" the temperature rises', and 'when the state is "off" the temperature falls'.

In a stochastic process the device-environment interactions are known only probabilistically. Inventory control may be viewed in this context. The warehouse (device) must be operated with minimum cost while keeping sufficient stock on hand (level of performance) to meet the problems of random demand and random time-of-delivery lag (environment). The plan's control is based on two criteria: (1) how large should an order be to replenish stock, and (2) when should an order be placed. According to the known probability distributions of the environment, the resultant fluctuating inventory level (process) can be predicted.

Adaptive plans must deal with incomplete information (deterministic or stochastic) about processes. Both the environment and the performance level of the device are to some extent unknown. By using information acquired during device-environment interaction, an adaptive plan reduces this uncertainty. If the distributions of the

demand and time-of-delivery lag are unknown in the inventory control problem discussed above, then the plan would have to be adaptive.

This thesis is concerned specifically with a special class of adaptive plans, called reproductive plans. Such plans will be fully and formally characterized later. (It might be noted that the terms "adaptive plan" and "reproductive plan" can be somewhat misleading. Since the plans do not adapt or reproduce but rather control adaptation and reproduction, better terms might be "adaptation plan" and "reproduction plan". In harmony with prevalent usage, however, the former terms will be used throughout this thesis.)

Adaptation appears in many natural situations such as evolution, or man-made instances such as artificial intelligence, control theory, and economic planning. Different situations have different devices to optimize, different methods to accomplish the necessary modifications, different environments to adapt to, and different performance criteria to attain. Because these instances are very diverse the terms "adaptive" and "adaptation" have been used with many different meanings in the literature. However these inconsistencies are usually not attributable to ambiguous definitions of these terms. In his review of these terms, Gaines has shown that most researchers accept the following as a definition of the behavioral aspects of adaptation.

Adaptation is a dynamic process which enables an "organism" to achieve and maintain a certain level of performance (sometimes termed fitness) when interacting with an environment. [8]

The terminological inconsistencies have arisen in the realization of definitions of adaptive behavior as implementations that behave adaptively. Theoretical unification thus requires an acceptable general paradigm, or functional definition, of adaptation. Recognizing this requirement, John H. Holland and his colleagues at the University of Michigan have begun to formalize adaptation. The theory is intended to be adequate for generating hypotheses about natural adaptation and for developing new algorithms (artificial adaptation). The framework is intended to be sufficiently encompassing for direct comparison of widely differing adaptive systems.

The function of an adaptive plan is to successively modify a device, producing a trajectory through the space of all possible devices. Due to incomplete information about the environment with which the device is interacting, it is uncertain whether a modification will be beneficial or deleterious. The essence of adaptation is the use of information received from the environment ("feedback") to modify the device. The feedback function can have many local optima. An efficient plan converges quickly to devices whose feedback corresponds to these optima. An effective plan's search prevents entrapment on

these early local optima by continuously locating better local optima in search of the global optimum.

One method that can be used to search the space is enumeration. A set of rules is used to generate every possible device and the "best-to-date" device is replaced only when a better one is generated. Assuming that the environment remains stable for the duration of the search, enumeration guarantees that the optimal device for any environment can be found. Enumeration is thus universally effective. However the search is almost always extremely inefficient because information received from testing a device does not affect the procedure that constructs the next device. Other methods frequently converge quickly to suboptimal devices.

Holland has developed the "reproductive plan" as a candidate for an adaptive plan which is both efficient and effective in a wide range of environments. Much of the development has been influenced by observing natural adaptive methods and generalizing them for use in any situation requiring adaptation. This special class of adaptive plans works in the following way. An initial set of devices is encoded into strings of attributes called structural representations, or loosely just structures. In each reproductive cycle, the current structures are tested against the environment to determine their level of performance. Structures are then reproduced (copied) a number of times according to their relative fitnesses.

Modification procedures (operators) are then employed at specified rates to produce new structures. In the type of reproductive plan of interest to Holland, the operators are generalizations of the genetic operators crossover, inversion, and (point) mutation. This thesis is concerned with using a reproductive meta-plan to study meta-adaptation of the genetic operator application rates and probabilities of crossover at successive locations in structural representations.

Chapter 2 of this thesis introduces Holland's theoretical framework, the basic elements being the environment, the adaptive system (structures, operators, feedback, and the plan), and the performance criterion. Chapter 2 concludes with a survey of some computer studies of reproductive plans which have uncovered many problems associated with putting theory into practice.

One of the problems is finding appropriate rates of operator application for an otherwise fully specified reproductive plan. Different environments may need different operator application rates for optimal adaptation of the devices. In two of the studies different stages of the adaptation sequence have required different optimal levels of operator application. Another problem is maintaining high genetic variance in a small population without unduly disturbing the efficiency of the reproductive plan. Most techniques used show a tradeoff between these two factors.



An investigation of adaptation of genetic operator values (parameters that control the rate of application of the genetic operators mentioned earlier) is discussed in Chapter 3. A detailed discussion of Cavicchio's parameter modification scheme (a non-reproductive meta-adaptive plan) reveals some problems to be solved. A description of a proposed reproductive meta-adaptive plan (henceforth called reproductive meta-plan or, when there is no confusion, meta-plan) concludes the chapter.

Chapter 4 discusses experimental results obtained from testing an implementation of the reproductive meta-plan proposed in Chapter 3 and an implementation of Cavicchio's parameter modification scheme, both applied to the same reproductive plan. The discussion focuses on: (1) the reproductive meta-plan's more vigorous search of the meta-space (combinations of genetic operator rates) enabling it to cope better with "crisis" situations than Cavicchio's meta-plan, (2) the lack of overall superiority of either meta-plan in non-"crisis" situations, and (3) the significance of biasing crossover and mutation location probabilities to promote more efficient adaptation by the adaptive plan.

Chapter 5 concludes this thesis with a discussion of further avenues for research specific to the reproductive meta-plan developed here and for research on reproductive plans in general.

## CHAPTER 2

### REPRODUCTIVE PLANS

Problems of optimization which are complex (optimum may never be found) and involve uncertainties (usefulness of available options is initially unknown) arise in fields such as genetics, artificial intelligence, game theory, systems control, economic planning, and psychology. Although some form of adaptation is a common solution to these diverse problems, the fundamental similarity of the solutions is often obscured by difficulties in identifying the underlying mechanisms of adaptation. It is often hard to determine (1) what is actually adapting, (2) what the mechanisms of adaptation are, and (3) what aspect of the environment is being adapted to.

The phenomenon of industrial melanism illustrates such difficulties. Moth species that were originally light coloured became dark after many years of industrial pollution had caused their camouflage, treebark, to turn dark. What is adapting? It is not the colour of the moth which is adapting, but rather the genetic material determining that colour. What are the mechanisms of adaptation? It is not natural selection (the payoff function), but rather genetic operators. It should be noted that the genetic operators do not start working when the change in

environment is noticed but are constantly operating, continually supplying new structures for natural selection to rank with respect to fitness. What aspect of the environment is being adapted to? It is not the increase in the amount of industrial pollution, nor the darkening of the tree bark. What is being adapted to is the ability of the moth's predator to see a light coloured moth against a dark background much more easily than a dark coloured moth. This seemingly subtle difference is important as it allows the feedback function (utility) to be used to describe the environment (see Section 2.1).

The fundamental questions about adaptation can also be applied to an artificial system, such as Samuel's Checker Player [11]. What is adapting? It is simply the coefficients of the evaluation polynomial. What are the mechanisms of adaptation? Operators which adjust the coefficients on the basis of feedback, as well as binary connecting operators which combine pairs of detectors. What is being adapted to? The Checker Player adjusts the coefficients to improve the accuracy of the evaluation function.

## 2.1 Holland's Framework

A unified theory of adaptation must be general enough to provide coherent explanations of both natural phenomena and algorithms used by artificial systems. Such a theory has been proposed by John H. Holland [9].

Explication of his theory begins with specific formal realizations of the general terms in the following (loose) definition of adaptation: some thing is, by some means, improving with respect to some extrinsic criterion.

The thing Holland calls a structure, denoted  $A$ . (More generally,  $A$  may be a collection of structures.) In ecology structures would correspond to organisms, in genetics, chromosomes, and in artificial intelligence, programs. The set of all possible structures is denoted  $\mathcal{A}$ . Since adaptation is a modification procedure, those tasks of interest have structures with identifiable substructural units, or more simply subunits. For example in Samuel's Checker Player [11] the structures would correspond to all  $k$ -tuples of weights, while the subunits would be the individual weights (signature tables in his later study [12]). For chromosomes the subunits would be genetic loci. Associated with each structure  $A \in \mathcal{A}$  is an information vector  $I_A \in \mathcal{I}$  (where  $\mathcal{I}$  is the set of all information vectors). The structures can be expanded to include an explicit finite record of past information vectors, or memory, denoted  $\mathcal{M}$ . Thus in general  $\mathcal{A} = \mathcal{A}_1 \times \mathcal{M}$ , where  $\mathcal{A}_1$  is the set of memoryless structures described above.

The extrinsic criterion for adaptation is the environment, denoted  $E$ . The environment consists of everything external to the adaptive system, such as the ecosystem of an organism or the data for a program. The

sequence of states of the environment may be governed by either a deterministic or a stochastic process. In either case the environment may also be (or not be) subject to sudden and possibly large modifications.  $E$  communicates the outcome of testing a structure  $A$  in  $E$  via an instance of the information vector denoted  $I_A^*$ . Thus an environment,  $E$ , can be viewed as a set  $I_E^*$  consisting of one  $I_A^*$  for each  $A \in \mathcal{A}$ . The initial uncertainty in the adaptation problem can be formalized as not knowing which  $E \in \mathcal{E}$  will confront the plan, or equivalently, not knowing which set of information vectors,  $I_E^*$ , will describe the outcomes of the tests. In Samuel's Checker Player this initial uncertainty can be viewed as not initially knowing which sequence of moves,  $E$ , from the set of all legal sequences of moves,  $\mathcal{E}$ , will face the Checker Player.

The means of adaptation, the procedure which modifies the structures(s), is the adaptive plan,  $\tau$ . A trajectory through the set  $\mathcal{A}$  is produced by successive generation of new structures. To be adaptive the trajectory must be influenced by the environmental input  $I_E^*$ . Hence if  $\mathcal{A}(t)$ , the set of structures tried at time  $t$ , is considered the state of the plan  $\tau$  at time  $t$ , then  $\tau$  can be considered a state transition function

$$\tau : I_E^* \times \mathcal{A} \rightarrow \mathcal{A},$$

that is  $\mathcal{A}(t+1) = \tau(I_E^*(t), \mathcal{A}(t))$ . However  $\tau$  can also be realized as a stochastic process

$$\tau : I_E^* \times \mathcal{A} \rightarrow \mathcal{P}$$

where  $\mathcal{P}$  is a set of probability distributions over  $\mathcal{A}$ . Thus  $\mathcal{A}(t+1)$  is chosen from  $\mathcal{A}$  according to the distribution  $\mathcal{P}(t+1) = \tau(I_E^*(t), \mathcal{A}(t))$ ,  $\mathcal{P}(t+1) \in \mathcal{P}$ . The deterministic case can be considered a single point distribution.

Since the structures are generated, some operator  $\omega \in \Omega$  is applied to  $\mathcal{A}(t)$  to give  $\mathcal{A}(t+1)$ . Thus  $\tau'$ , the underlying operation of the plan  $\tau$ , can be represented as a function

$$\tau' : I_E^* \times \mathcal{A} \rightarrow \Omega$$

where  $\Omega = \{\omega : \mathcal{A} \rightarrow \mathcal{P}\}$ . Examples of operators are learning rules in artificial intelligence and mutation and crossover in genetics. Once the operators in  $\Omega$  are specified,  $\tau'$  determines  $\tau$ .

The improvement of the structure(s) corresponds to an increase in the fitness (utility, payoff),  $\mu_E$ , of the structure(s) in the particular environment  $E$ .  $\mu_E$  is a function

$$\mu_E : \mathcal{A} \rightarrow \text{Reals}$$

For many tasks,  $\mu_E$  can be multidimensional and, in any or all dimensions, nonlinear, discontinuous, and multimodal, creating intriguing problems that cannot be solved by analytic techniques.  $\mu_E$  is usually one of the components of  $I_A \in \mathcal{I}$  and in the cases where  $\mu_E(A) = I_A^*$  the plan is termed payoff only. Improvement can be defined as

$$\mu_E(\mathcal{A}(t+n(t))) \geq \mu_E(\mathcal{A}(t)),$$

for all  $t > 0$ , and some  $n(t) > 0$ . (In those tasks that

require monotonic improvement  $n(t) = 1$ .)

The foregoing features can completely specify an adaptive system,  $\langle A, \Omega, \mathcal{I}, \tau \rangle$ . However when studying adaptation it is often necessary to compare the efficiency of all the feasible plans for a task,  $\{\tau_1, \tau_2, \dots, \tau_n\} = \mathcal{T}$ , under the uncertainty represented by  $\mathcal{E}$ . The criterion,  $\chi$ , used to compare the plans in  $\mathcal{T}$  often depends on the particular task, but usually involves accumulation of payoff. Thus a problem in adaptation is said to be well posed when  $\mathcal{T}$ ,  $\mathcal{E}$ , and  $\chi$  have been specified for a given adaptive system.

## 2.2 Reproductive Plans

In addition to providing a general framework for investigating adaptive processes, Holland has also proposed a class of adaptive plans, called reproductive plans, which he has shown to be robust. Intuitively, a robust plan is efficient in all environments. More formally, robustness is associated with optimal allocation of trials to structures for all feedback functions.

The search of the set  $\mathcal{A}$  must continue as long as there are significant improvements to be made. This search however poses two conflicting methods of finding better structures, (1) exploitation of known subunits in new contexts, and (2) search for alternative subunits that may lead to optimal performance. The adaptive plan must therefore continue to test and incorporate new and old

subunits associated with better performance. To identify useful subunits, structures in  $\mathcal{A}$  must be compared. Comparing common subunits leads to the problem of apportioning credit to those thought to improve performance. Problems of apportioning credit arise when performances of subunits combine non-linearly. Thus a method for associating combinations of subunits is necessary. Holland uses the notion of a schema, denoted  $\xi$ , for this purpose. Holland has shown that reproductive plans search the set of schemata, denoted  $\Xi$ , instead of  $\mathcal{A}$ , allowing intrinsic parallelism in the plan's operation.

### 2.2.1 Schemata

One class of subunits, used to represent a structure in a form suitable for use by a reproductive plan, is a set of detectors,  $\{\delta_i: \mathcal{A} \rightarrow V_i, i=1, \dots, l\}$ . Each  $\delta_i$  maps an effectively describable feature of a structure into a set of values (attributes) for that feature. Thus the representation of a structure  $A \in \mathcal{A}$  would be  $(\delta_1(A), \delta_2(A), \dots, \delta_l(A))$ . Henceforth the set  $\mathcal{A}$  and the associated representations will be considered equivalent.

A schema,  $\xi$ , is a subset of  $\mathcal{A}$  having some common attributes. Certain attributes are fixed while the remainder may vary independently (such an attribute is symbolised  $\square$ ). The set of schemata,  $\Xi = \prod_{i=1}^l [V_i \cup \{\square\}]$  decomposes  $\mathcal{A}$  into subsets (if  $k$  positions are fixed the schema could also be considered as a  $1-k$  dimensional

hyperplane in  $\mathcal{A}$ . A schema is said to be defined on the positions that are not  $m$ 's. Hence all schemata defined on the same positions partition the set  $\mathcal{A}$ . From the other point of view a binary structure represents  $2^l$  schemata. The length of a schema,  $l(\xi)$ , is defined to be the distance between the two end defining positions. For example, the length of the schema  $\square\square\square_{x_1}\square\square\square_{x_2}\square\square\dots\square\square_{x_k}\square\square$  is  $(x_k - x_1)$ .

Associated with each schema  $\xi$  there is a utility  $\mu_\xi$  equal to the average of all  $\mu(A_i)$  for all  $A_i \in \mathcal{A}$  that are instances of  $\xi$ . For each observation  $A_i \in \xi$ ,  $\xi$  is accredited the utility  $\mu(A_i)$  (this is done for all  $\xi$  of which  $A_i$  is an instance). Given a set of observations  $\mathcal{A}(1), \mathcal{A}(2), \dots, \mathcal{A}(t)$ , the observed average payoff  $\hat{\mu}_\xi$  of the observed instances of  $\mathcal{A}(i) \in \xi$  becomes a better estimate of  $\mu_\xi$  as  $t$  (time) increases.

### 2.2.2 General Scheme of Plans of Type $\mathcal{R}$

In general a reproductive plan allocates tests to structures according to their relative fitness in the environment. More specifically, given a finite sample of structures  $\mathcal{B}(t) = \{A_1(t), A_2(t), \dots, A_M(t)\} \subset \mathcal{A}$ , the population at time  $t$ , the following is a basic algorithm for reproductive plans (henceforth called plans of type  $\mathcal{R}$ ):

- (1) Select one structure from the population at time  $t$  probabilistically, after assigning each structure a probability proportional to its observed performance at time  $t$ .

(2) Copy the structure and apply operators to produce a new structure.

(3) Either (3.1) or (3.2).

(3.1) Select another element of the population (all elements equally likely) and replace it with the new structure.

(3.2) Store the new element returning to (1) until the number of stored elements equals  $M$  and replace the whole population with the new structures.

(4) Observe and record the performance of the new structure(s).

(5) Increment  $t$  and return to (1).

In the above algorithm it can be seen that the formal framework discussed in Section 2.1 has been satisfied.  $\mathcal{A}$  can be any set of structures with string representations, if generalized genetic operators are used (see Section 2.2.3).

$\mathcal{M}$  is simply the retained performances,  $\mu_E(\mathcal{A}_h(t))$ ,  $h=1, \dots, M$  of the structures in the population. The information,  $\mathcal{I}(t)$ , is the payoff of the new structure(s).

The operators are of the form  $\omega : \mathcal{A} \rightarrow \mathcal{P}$ . Thus  $\tau$ , the plan described above, has the (correct) form  $\tau : I_E^* \times \mathcal{A} \rightarrow \mathcal{P}$ .

### 2.2.3 Generalized Genetic Operators

Although the operators in a general reproductive plan can be any which satisfy the criterion of the

framework, genetic operators confer two advantages:

- (1) intrinsic parallelism, in that a great number of schemata are affected, and
- (2) compact storage and use of information resulting from previous observations of schemata, i.e. if schemata are ranked by their relative frequency of representation in  $\mathcal{B}(t)$ , then genetic operators exploit this information without much disturbing the ranking of short schemata, because most of the short schemata in a structure are not affected by a single application of a genetic operator.

Three operators are necessary and sufficient to generate any structure in  $\mathcal{A}$ : crossover, inversion, and mutation.

The following discussion introduces basic concepts from genetics essential to a full appreciation of the role of genetic operators in a reproductive plan. A gene is a functional unit, the counterpart of parameters or detectors. An allele is a particular instance of a gene, analogous to a value for a parameter or a detector and in most cases is in a range of acceptable values. A chromosome is a string of gene loci. (For present purposes all the genes may be regarded as placed on a single chromosome, as is true in bacteria but not in higher organisms.) An individual in a population can be haploid (1 chromosome), diploid (2 homologous chromosomes), or polyploid (more than 2 chromosomes). If an individual is haploid the terms

individual and chromosome will be used interchangeably. A locus is a physical position on the chromosome. Loci which are closer together have a greater degree of linkage. A non-haploid individual may have different alleles at homologous loci. In such cases one allele can dominate the others functionally. The particular allele that gives the locus its value is dominant; the others are recessive.

Crossover serves to recombine alleles by exchanging segments between pairs of chromosomes. This operator works in three steps: (1) two structures,  $A = a_1 a_2 \dots a_l$  and  $A' = a'_1 a'_2 \dots a'_l$ , are randomly chosen from the current population (in reproductive plans the choice is not uniformly random for all structures but is proportional to utility), (2) a breakpoint,  $p$ , is randomly selected, and (3) two new structures are formed by exchanging the segments to the right of the breakpoint. The two resulting structures are  $a_1 a_2 \dots a_{p-1} a'_p \dots a'_l$  and  $a'_1 a'_2 \dots a'_{p-1} a_p \dots a_l$ . This operator affects schemata in two basic ways. First, new instances (different contexts) of schemata already in the population are generated. Second, new schemata are generated. The two effects do not significantly disturb the ranking process except for longer schemata. Thus a linkage phenomenon is induced (shorter schemata are less affected than longer ones). Each crossover event affects a large number of schemata (intrinsic parallelism) and each schema is affected independently of what happens to other schemata, since the probability of  $\xi$  being in the population at time

$t+1$ ,  $P(\xi, t+1)$ , depends only on  $l(\xi)$ ,  $\hat{\mu}_\xi$ , and  $P(\xi, t)$ .

Inversion alters the linkage between alleles by changing the length of schemata. The three steps of this operator are: (1) a structure  $A = a_1 a_2 \dots a_l$  is selected from  $\mathcal{B}(t)$ , (2) two breakpoints,  $x_1 < x_2$ , are randomly selected, and (3) a new structure is formed by inverting the segment between the two breakpoints resulting in the structure  $a_1 a_2 \dots a_{x_2} a_{x_2-1} \dots a_{x_1+1} a_{x_1} \dots a_l$ . This new structure (and the affected schemata) differs from the original one only in the order of the loci that the genes occupy. Under plans of type  $\mathcal{P}$  instances of shorter schemata proliferate more rapidly since they are less likely to be removed from the population by crossover. Inversion alone is sufficient to produce all permutations of a structure. Like crossover, inversion is intrinsically parallel.

Mutation is the random replacement of alleles to produce a new structure. Generally the probability of mutation at a locus is small (usually less than the inverse of the population size, though not dependent on it). This operator acts upon the structure  $A = a_1 a_2 \dots a_l$  in two steps: (1) loci  $x_1, \dots, x_h$  are determined to undergo mutation, then (2) a new structure  $A' = a_1 \dots a'_{x_1} \dots a'_{x_1} \dots a'_{x_h} \dots a_l$  is formed where  $a'_{x_i}$  is randomly chosen from the range  $V_{x_i}$  of  $\delta_{x_i}$ . Mutation serves two roles. First, it guarantees a finite expected interval of time between the loss and reappearance in the population of an allele at any locus. Second, in those cases where

genes have many alleles it serves as an enumerative process, that is, it can search small areas near the unmutated structure. Mutation is also a constant source of loss for schema  $\xi$ . Thus it can be viewed as a disturbance to prevent entrapment on local optima.

In addition to these three operators, which are necessary and sufficient to produce any structure in  $\mathcal{A}$ , Holland discusses the potential use of some operators analogous to other genetic mechanisms such as dominance. Addition of such operators might improve a plan's efficiency.

### 2.3 Optimal Allocation of Trials and the Robustness of Reproductive Plans

The observed average performance of instances of  $\xi$ ,  $\hat{\mu}_\xi$ , is a sample average and is therefore subject to sampling error. A conflict thus arises. Should an adaptive plan exploit this knowledge by allocating trials to the observed best? Or should it obtain more information by allocating trials to other than the observed best to reduce the probability of error? Each alternative is a source of loss of performance. If a trial is allocated to the observed best it may incur a loss because the observed best is not the best. Likewise if a trial is allocated to a schema that is not the observed best a loss is incurred if the observed best is the best among known alternatives.

Holland has claimed that for convergence to

minimal expected loss (a measure discussed further below) an adaptive plan must allocate to the observed best schema a number of trials which is an exponential function of the number allocated to the remainder. The proof of this result (Theorem 5.3 [9]) actually holds only for normally distributed schemata. But simulations by Prantz (unpublished) suggest the same is true for other distributions.

The concept of robustness needs a precise definition if any mathematical results are to be obtained. Convergence of a plan to an optimal structure is an immediate candidate. But enumerative plans converge. And the observed utility of a suboptimal schema has a non-zero probability of being greater than that of an optimal schema. Repeated testing of the structures containing these schema will cause optimal structures to be displaced with a limiting frequency approaching that probability. These difficulties with the convergence measure have led Holland to choose the minimum loss criterion as the test for robustness.

Holland has shown that plans of type  $\mathcal{P}_1(P_C, P_I, P_M, \langle c_t \rangle)$  are robust. This special reproductive plan uses the three genetic operators crossover, inversion, and mutation at a locus.  $P_C, P_I, P_M$  are the rates for the respective operators.  $\langle c_t \rangle$  is an infinite sequence that converges to 0 and whose infinite series is unbounded. This sequence is used only to obtain some of the mathematical

results about plans of type  $R_1$  and is not important in any of the following discussion.

#### 2.4 Bagley's Results

Bagley [1] was the first to implement a reproductive plan based on (an early version of) Holland's theory. The main aim of his research was the comparison of a reproductive plan with a correlation algorithm. A correlation algorithm is a prototype of many weight adjusting techniques such as the Uhr Vossler pattern recognition system [14], and Samuel's Checker Player [11],[12]. A correlation algorithm is an adaptive plan which determines the worth of parameter values (or parameter value pairs, triples, etc.) according to encounters of parameter value vectors (structures) with the environment.

Bagley chose a game called hexapawn to test the two algorithms. The hexapawn player used a strategy based on the present board position and the value of the particular parameter in the parameter value vector that controlled the next move from that board position. Each environment,  $E$ , was a fixed playing strategy. The adaptive plan then had the task of generating the parameter values (structure) that would have obtained a win. The restriction of the opponent's strategy to a particular fixed strategy enabled Bagley to choose environments with predefined depths. The depth corresponded to the cardinality of the largest set of interacting parameters.

The experience array, which contained utilities of the values for each parameter (pair, triple, etc.) could be composed of single parameters and their possible values, parameter pairs and their possible value combinations, etc. The size of the set of interacting parameters was called the level of the adaptive plan. (The term "level" is used here to remain consistent with Bagley's terminology. It should not be confused with the use of the term "second level" which is used to describe the reproductive meta-plan.) Note the similarity of the level of a correlation algorithm and signature tables in [12]. This ability to control the depth and level was the reason for choosing hexapawn. By calculating the number of possible parameter value vectors that could have been generated to find the maximal vector Bagley showed that the correlation algorithm with unmatched (unequal) level and depth did much worse than the matched version, and that levels greater than three were almost unmanageable.

The "chromosomes" used by Bagley's reproductive plan were strings of parameter values. Each parameter was included once in the chromosome. The reproductive plan used large populations of diploid individuals (approximately 200) that incorporated pure dominance. Crossover, inversion, and mutation were the genetic operators. Bagley demonstrated that the reproductive plan outperformed the unmatched correlation algorithm and in some cases the matched correlation algorithm.

In addition to this important result Bagley discussed two interesting ideas: (1) the use of meta-adaptive plans that would control values of parameters used by the adaptive plan, and (2) the use of genes to control breakpoints and mutation locations. Both of these concepts will be considered in subsequent chapters.

### 2.5 Cavicchio's Results

Cavicchio's research [5] was concerned with developing efficient and effective reproductive plans. One of the methods developed was a meta-adaptive plan that adjusted the genetic operator rates during adaptation. The system that Cavicchio made adaptive was the pattern recognition scheme of Bledsoe and Browning [4]. The environment was composed of two 16-letter alphabets which were represented by a 1 in every grid point that the letter touched, and 0 elsewhere. The system "learned" one of the alphabets (each letter of the alphabet was accepted as an example of that letter and was used for comparison purposes) and then tried to correctly match the letters of the other alphabet with this "learned" alphabet. The haploid individuals were strings of detectors ( $n$ -tuples ( $n=2,3,\dots,6$ ) of grid points). Each individual in the initial population had the same number of grid points randomly distributed in a variable number of genes (detectors). The pattern recognizer used the detectors to classify unknown letters as examples of known letters

according to the number of matches made. To achieve a match between an unknown letter and a known letter for a given n-tuple the two letters had to coincide exactly on those grid points represented by the n-tuple.

The plan searched  $\mathcal{A}$  to find a structure that could have correctly matched the unknown letters with the known ones. Each individual's utility was measured by its ability to correctly classify the unknown letters (actually a function of its ability to discriminate between the two likeliest candidates). The following procedure was used to calculate each structure's utility. The number of matches between each unknown letter and each known letter was calculated as a percentage. The two highest percentages for each unknown letter were chosen to designate the two likeliest classes to which the unknown letter belonged. If the unknown letter belonged to neither class the score,  $s$ , was 0. If the unknown letter belonged to one of these two classes,  $s$  was a function of a discrimination value,  $d$ :

$$s = \begin{cases} 100 & \text{if } d \geq 10 \\ 5d + 50 & \text{if } -10 < d < 10 \\ 0 & \text{if } d \leq -10, \end{cases}$$

where  $d$  was the percentage value of the class to which the letter belonged less the value of the other class. The utility for a device was the average of the scores for all unknown patterns. An optimal structure would correctly classify the letters in the second alphabet.

Each generation of Cavicchio's adaptive plan can

be divided into the following phases: (1) preselection of offspring (from the previous generation) and selection of parents based on utility, (2) readjustment of genetic operator values using the meta-plan, (3) reproduction of the parents and use of the operators crossover, double crossover (interchange of two segments), inversion, and mutation to form the offspring, and (4) assignment of a utility to the offspring according to their fitness in the environment.

During reproduction the parents were copied (reproduced) according to their reproduction quotas. When the required number of new population members had been generated, the genetic operator rates were modified. The copies of the parents were then modified by the genetic operators.

Crossover and double crossover could not happen concurrently. A method was devised to allow only one of these operators to be used. The two crossover rates were summed to provide a new rate of crossover. Then using the individual rates as a two value discrete probability distribution one of the crossover operators was chosen if a crossover was to occur.

Finally it was decided whether each of the mutation operators was to be applied. Mutation 1 replaced the complete n-tuple with n randomly chosen grid points. Mutation 2 changed one grid point in the n-tuple. Mutation 3 changed two genes into one or one gene into two depending

on the number of grid points found in the gene to be mutated. If a mutation operator was used, the number of mutations was a choice of an integer in the range from 1 to  $\lfloor \text{value} * 5 \rfloor$ , where  $\lfloor x \rfloor$  was the largest integer  $\leq x$ . The offspring were then assigned a utility.

Experimentally, Cavicchio developed some techniques to increase the effectiveness of adaptive plans. One of the primary causes of loss of a continual effective search was a decrease in population variance, because different individuals came to contain a large proportion of common genes. This problem was especially prevalent in the small (12 to 40 individual) populations Cavicchio used. A selection scheme which limited the number of offspring each parent was capable of producing was developed to prohibit a good individual from reproducing so many offspring that the population tended to resemble it. An M/N population specified the maximum size of the population, M, the number of individuals kept as potential parents, N, and the number of offspring produced at each generation,  $M - N$ . The N best individuals were the potential parents. Those individuals that were not in this group in the previous generation were called new population members. The selection scheme chose parents from these N individuals (based on utility) and granted a reproduction quota to each parent (based on relative utility) such that the sum of the individual quotas was  $M - N$ . Each parent could reproduce a maximum of two offspring per generation and a maximum of

three offspring over all generations. The latter condition was relaxed whenever the sum of the reproduction quotas was less than  $M - N$  and all parents had been allotted their maximum quotas. In this case parents were chosen at random until the offspring quotas reached the desired total.

Cavicchio also developed a preselection scheme that prohibited an offspring from entering the current population (those individuals that reproduced in the next generation) unless it was superior to its parent (if formed by a mutation operator) or one of its parents (if formed by a crossover operator), even though it may have been better than some other member of the current population. If the offspring became a new population member, the inferior parent was removed from the current population.

To increase the plan's efficiency Cavicchio developed a number of techniques of which two worked: chromosomal extension using intrachromosomal duplication, and mutation pools. The extension method started with short chromosomes. At certain times these chromosomes were extended by duplicating parts of themselves until the maximum length was reached. This method was intended to make maximum use of the available detectors before extension took place. The plan tended to find peaks quickly and then expanded the search space leaving these peaks intact. The use of mutation pools made mutation less random. Whenever mutation was to take place a gene from a good individual was used to replace the mutated gene in the

offspring individual. The mutation pools were updated at specific times.

To increase both efficiency and effectiveness Cavicchio developed a parameter modification scheme which will be discussed more thoroughly in Section 3.2. This scheme altered the rates of involvement of the genetic operators (crossover, mutation, and inversion) in creating new individuals. It was intended to increase efficiency by allowing more large steps (crossover) in the space *when* needed, and to increase the effectiveness by increasing the number of small steps (mutation) and decreasing the number of large steps when needed.

Cavicchio's very favorable results showed a marked improvement over the individual weight adjustment technique (a correlation algorithm) used by Uhr and Vossler [14]. Aspects of Cavicchio's reproductive plan will be investigated in more detail in Chapters 3 and 4.

## 2.6 Hollstien's Results

Hollstien [10] was interested in developing evolutionary techniques for computer control systems. Thus his plans were typified by very fast convergence (20 generations) to a structure which was intended to maximize a function of two variables. In order to get this desired speed he modelled the plan on artificial breeding techniques commonly used in plant and animal breeding.

These breeding techniques have two distinct

phases: (1) selection of parents and (2) mating for reproduction. The selection methods that Hollstien considered were (1) progeny testing (only parents that produce desirable offspring were allowed to continue reproducing), (2) individual selection (parents were selected on their own merit), (3) family selection (all members of the family with highest mean value were chosen as parents), (4) within-family selection (the most valuable members of each family were chosen), and (5) combined selection (two or more of the preceding methods were used concurrently). He considered the following mating techniques: (1) random mating, (2) inbreeding (mating of related parents), (3) linebreeding (mating all parents to one other parent), (4) outbreeding (mating of parents that are less closely related than random mating would have caused), (5) self-fertilization, and (6) clonal propagation (duplication of an individual).

The set  $\mathcal{E}$  consisted of fourteen mathematical functions. Each function had peculiar search problems associated with its particular characteristics. To develop a robust plan Hollstien tested different breeding techniques using populations of 16 diploid individuals with binary and Gray code representations of the two function variables. The plan used crossover and mutation as the genetic operators. The experimental results established that (1) the Gray code representation worked better due to the adjacency properties of the code, (2) genetic drift was

severe in small populations and had to be combatted by the adaptive plan, (3) a recurrent breeding plan that used inbreeding and crossbreeding was effective for the functions tested, (4) different types of dominance benefited the overall performance of the system, (associating a modifier gene that controlled dominance with each gene allowed evolution of dominance also), (5) a reproductive plan induced a more robust and more efficient adaptive search than either global or local random searches, and (6) the selection of parents based on individual utility was ineffective when applied to multi-modal functions.

## 2.7 Frantz' Results

Frantz investigated some of the underlying functions of reproductive plans. He initiated his experiments with the intention of showing that reproductive plans adapt differently to non-linear environments than to linear ones, and that resulting differences in adaptation can be detected and used to provide more information about the environment. The two effects that Frantz studied were position effect and frequency of combination effect.

Frantz called genes which combined non-linearly, in a particular environment, dependent genes. The defining distance of dependent genes is similar to the length of a schema, i.e. the rightmost dependent gene locus minus the leftmost dependent gene locus.

Position effect was Frantz' term for the effect

that the defining "distance of dependent genes has on adaptation. The effect has been observed for pairs of genes in a population at equilibrium, when the level of adaptation is such that no significant further evolution of the population as a whole is taking place.

In the frequency of combination effect, the frequency of a highly ranked group of alleles increases faster than would be expected according to the frequency of the individual alleles defining the group. This effect can easily be seen in a two gene model but is difficult to extend to a larger, more complex model.

Frantz' reproductive plan worked with a haploid population of size 100. Crossover, inversion, and mutation were used as genetic operators. This was the first study that followed the latest version of Holland's theory exactly.

Frantz' experiments that dealt with position effect were of two distinct types. First, he wanted to test the hypothesis that the defining distance of the dependent genes affected the average payoff of the population. He found no significant difference in equilibrium populations. However the work with evolving populations showed some significance in early generations in difficult environments, which he concluded meant position had an effect if the original population started far from the optimum and the environment had many false peaks. Second, his experiments that dealt with inversion (to achieve close permutations of

dependent genes) showed no significant change in distance. He concluded that short chromosomes and the small number of generations prevented this operator from accomplishing its intended purpose. (One thing that he did not mention is that if the operator did work in later generations it would have been hard to see since the adaptive advantage of position was not significant in equilibrium populations as was pointed out earlier.)

In experiments on frequency of combinations Frantz used a chi-squared contingency test to detect non-linearities between genes. These experiments also helped to confirm the hypothesized position effect.

Frantz discussed a practical use for these intrinsically important results. The probability of a reproductive plan finding the true optimum of a function can be increased by devising good permutations of dependent genes on the chromosomes. Using a chi-squared contingency test analysis of preliminary experiments, an experimenter can discover, with a high probability of success, groups of dependent genes. Grouping these dependent genes closely on the chromosome gives the desirable permutations.

## 2.8 De Jong's Results

De Jong was concerned with the design and analysis of adaptive systems especially adaptive computer software. He was particularly interested in the evaluation of robustness. He chose two performance criteria. Online

performance evaluated every response of the adaptive system, reflecting dynamic control of a computer system. Offline performance evaluated only those responses that improved system performance which reflected situations in which testing was done independently of the system being controlled.

De Jong's results encompass the most complete analysis of reproductive plans to date. To test the robustness of the plans the environments consisted of continuous, discontinuous, convex, non-convex, unimodal, multimodal, low dimensional, and high-dimensional functions.

De Jong's basic reproductive plan, R1, used two genetic operators, crossover and mutation. He was particularly concerned with allele loss (premature convergence) in his populations. In the theory Holland was not concerned with the problems of small populations. He varied four parameters: population size, crossover rate, mutation rate, and generation gap. The different parameter settings usually created a tradeoff between decreasing the allele loss rate and initial performance degradation. Some parameter settings created different performances in online and offline evaluations.

Generation gap was a parameter originated by De Jong. The gap is a value  $x$ ,  $0 < x \leq 1$ . The population reproduces offspring in the manner discussed in Section 2.2.2 (step 3.2). When the number of offspring equals that portion of the population indicated by the gap value, the

offspring replace randomly chosen parents and a new generation begins.

Recognizing basic limitations in plan R1, De Jong evolved five variations in an attempt to reduce the allele loss rate and increase both online and offline performance. The elitist plan, R2, ensured that the best individual to date was kept in the population. A good individual would less often be lost before its true importance was realized. The result was a more conservative sampling policy (similar to reducing the crossover rate which also increased performance) which caused a lowering of the allele loss rate and an improvement of both performance criteria except on highly multimodal functions.

Stochastic side effects occurred in these plans for two reasons. First the error involved in the sample means affected the selection probabilities. The sample means were made more reliable by increasing the population size with the attendant loss of early performance gains. The second source of error resulted from taking a finite sample from  $\mathcal{B}(t)$  according to a selection distribution. Thus the actual number of offspring produced by an individual could differ markedly from the expected number. Therefore the expected value plan, R3, forced the actual number of offspring to be within  $\pm 1$  of the expected value. This method was also used by Cavicchio. R3 performed significantly better than R1 but not as well as R2.

The logical next step was to form an elitist

expected value plan, R4. R4 performed better than any of the other plans so far described.

Even though R4 performed well it still had some problems finding the optimum on highly multi-modal functions. The problem of premature convergence (allele loss) had not been totally eliminated. Some available methods of lowering the allele loss rate were to increase the population size, the generation gap, and the mutation rate. In all these solutions the length of time required for a schema to saturate the population increased. Thus a suboptimal schema which appeared to be better than the rest of the population would take longer to saturate. Saturation was due to the use of small finite populations. The solution permitted exponential growth of the observed best without a rapid saturation.

The method used was an analogue of crowding in nature. Crowding occurs when large numbers of like organisms compete for finite resources, resulting in lower life expectancies and lower birth rates. In the theory, life expectancies can be considered as the expected length of time to replacement by an offspring. (The connection between life expectancy and generation gap was very significant in the experimental results.) Likeness was defined by De Jong in terms of a measure of matching alleles.

Based on these observations, De Jong developed the crowding factor plan, R5. His results showed a significant

performance improvement in the highly multimodal function and only slight degradation of performance in the unimodal functions.

De Jong's final effort, a generalized crossover plan, R6, tried to deal with the representational dependencies of his plans which occurred because his plans did not use the inversion operator. The results however were the complete opposite of what he expected. The allele loss rate increased and the performance degraded as the number of crossover points increased.

De Jong concluded his work by comparing the reproductive plans with a conjugate direction optimization technique and a variable metric method, two standard methods of function optimization. These methods were found to be much better for the functions for which they were developed, and much worse for the others.

## CHAPTER 3

### A REPRODUCTIVE META-PLAN

Holland has formalized the functional aspects of adaptation. Generalizing a well tested adaptive plan, natural evolution, he has shown that reproductive plans are robust. Arguing that adaptation is basically the same in whatever context it appears, Holland proposes that reproductive plans provide a general robust adaptive strategy for any structures representable as strings and any utility function.

Holland's colleagues and students have undertaken a number of computer simulation studies, which largely bear out the asserted robustness of adaptive plans. These studies have also revealed a number of difficulties associated with small populations of reproducing individuals. In the remainder of this thesis second-level reproductive adaptation will be considered as a possible improvement in reproductive plans. Even though the following discussion deals with the control of genetic operator rates, meta-adaptive plans are not constrained to this function. Other meta-adaptive procedures could involve controlling migration between subpopulations or adaptation of structural representations. (The latter method is actually discussed by Holland in Chapter 8 of [9].) Also,

it is not required that meta-adaptation be implemented as a second level plan. A more realistic analogue of natural systems would require that these control functions be part of the lower level structures. Genetic loci themselves would control the operator rates at other loci. It is doubtful that this alternative method would yield quick responses to environmental changes comparable to those of the present approach (see Section 4.2).

In what follows it will be necessary to differentiate between adaptive and meta-adaptive functions. Hence whenever confusion might arise any symbol associated with adaptation will be superscripted with a 0, whereas a superscript 1 will be associated with meta-adaptation. Also, terms like "meta-structures", "meta-parents", "meta-generations", and "meta-environment" will denote, respectively, structures, parents, generations, and environment employed in the meta-plan's genetic algorithm.

### 3.1 Meta-adaptation of Genetic Operator Rates

Each generation of a reproductive plan can be divided into several different phases, including the reproductive phase during which new structures are created, the evaluation phase in which the structures are assigned a utility after interacting with the environment, and the selection phase during which the next set of "parents" is chosen. This research is concerned with the reproductive phase.

The reproductive phase is that part of the reproductive plan during which the function  $\tau^0$  uses the set of "parent" structures to produce "offspring". In Cavicchio's implementation [5] the "parents" are copied (reproduced) and then crossover, double crossover, and inversion are applied to generate new individuals. The probability of a break occurring is the same for all gene boundaries. Three different types of mutation are used to modify the genes (see Section 2.5). Each gene has equal probability of being mutated by any given method.

During the reproductive phase each genetic operator may potentially be involved in the creation of each new individual. The adaptive plan determines the participation of each operator according to rates of certain parameters. Each of these values, called genetic operator rates, is the probability of using a specific genetic operator during the creation of a new individual. Although concurrent involvement of certain operators may be restricted, such restrictions are not relevant to this discussion.

During adaptation a reproductive plan can use a fixed set of genetic operator rates or a sequence of (sets of) genetic operator rates. A scheme that generates such a sequence of genetic operator rates is a meta-adaptive plan (or a "meta-plan"). It searches the meta-space  $\mathcal{A}^1$  to find the meta-structure (a string of genetic operator rates) which interacts favorably with the meta-environment,  $E^1$ ,

(the adaptive plan).

Investigators of reproductive plans have discovered some of the problems encountered with poorly chosen genetic operator rates. Bethke et al. [3] have shown a plan's sensitivity to operator rates by studying two indicators of adaptive ability: (1) the rate of convergence to an optimal structure (efficiency of the search), and (2) the rate of decay of population variance (as the population decreases in variability, efficiency and effectiveness also decrease). Experimental results showed a tradeoff between the factors that control these indicators. For example, high mutation rates at the outset produced rapid convergence and kept the population variability high. However as the population neared the optimal structure high mutation rates slowed convergence drastically, but still managed to keep the variability high. Thus settings which could be considered optimal for the duration of the adaptive period must sacrifice certain desirable features during different stages of adaptation. The most notable of these tradeoffs is between efficiency and effectiveness.

De Jong's results [6] on offline performance confirm these results. However online performance was degraded with high mutation rates even in the initial stages.

The possibility of different optimal rates at various stages of adaptation was considered by Cavicchio [5] when he encountered a similar tradeoff problem during his

research. Crossover could be considered as a large step in the search space, double crossover as a medium sized step, and mutation as a small step. He concluded that the rate of the crossover operators should have been decreased as the plan converged towards an optimal structure. This reduction should allow a larger proportion of the offspring to be modified solely by the mutation operator.

Thus it appears that even though there are sets of genetic operator rates that are optimal for the entire adaptive period, finding one of these sets would not be worthwhile for two reasons. First, every environment would require a different set of optimal genetic operator rates. Density of good structures, modality of the utility function,  $\mu^0$ , and other similar aspects of the space  $\mathcal{A}^0$  cause special search problems. Thus different  $\mathcal{A}^0 - \mathcal{E}^0$  combinations require vastly different methods (strings of genetic operator rates) for searching  $\mathcal{A}^0$ . It would be too costly to find an optimal set for each new situation and would be impossible if the environment changed. Second, the plan that uses the optimal set of genetic operator rates is not as efficient or as effective as the same plan with modifiable genetic operator rates, since the best combination of genetic operator rates for any particular generation may not be the set chosen as optimal for the whole period. The possibility exists that the modifiable plan could be better overall even if it did not produce currently optimal genetic operator rates at each

modification step.

A reproductive plan that is unable to modify its genetic operator rates is constrained by the initial settings. It is unable to vacillate between searching large areas of  $\mathcal{A}^0$  quickly and smaller regions in greater detail. Thus flexibility of genetic operator rates may provide some of the efficiency and effectiveness that is essential for a plan to be general.

### 3.2 A Critique of Cavicchio's Meta-adaptive Plan

Among the investigations reviewed in Chapter 2, Cavicchio's [5] was the only study of meta-adaptive plans. While investigating reproductive plans for pattern recognition (see Section 2.5) he noticed the generation of new population members (those individuals that become "parents" in the following generation) fell off after the population of structures,  $\mathcal{B}^0$ , had reached a certain performance level. Cavicchio hypothesized that the crossover application rate which had promoted an efficient search during the early stages of adaptation had become detrimental to the effectiveness of the search in the later stages.

Cavicchio therefore developed a meta-adaptive system (parameter modification scheme) for the following two reasons: (1) to increase the efficiency and effectiveness of the search at various stages of adaptation, and (2) to make the plan more general by not relying on the initial

genetic operator rates. He has succeeded to some degree in attaining these goals. Some problems remain.

Cavicchio's parameter modification scheme is a function which independently adjusts each current genetic operator rate. The modification depends on the observed frequency of the operator's involvement in creating new population members. The function is:

$$P(t+n) = P(t) + [P(t) - O(t+n)]A,$$

where:

$P(t)$  is the genetic operator rate at time  $t$ ;

$O(t+n)$  is the observed frequency of the particular operator's involvement in creating new population members between times  $t$  and  $t+n$ ;

$n$  is determined by requiring a certain number of new population members to be formed between modifications;

and  $A$  is a term ( $0 \leq A \leq 1$ ) which assures that  $m \leq P(t+n) \leq 1$ , where  $m$  is the minimum rate allowed for the particular genetic operator, and which dampens the modification. (For a more detailed description see [5] pp. 146-154.)

Other modifications were used by the meta-plan whenever the adaptive plan began to stagnate. Two such adjustments were: (1) halving the number of new population members required for modification to occur, and (2)  $P(t) \leftarrow (P(t) + m)/2$ , i.e. the genetic operator rate was set to half the distance between its present rate and

the particular operator's minimum rate.

Cavicchio achieved significant gains by incorporating the above parameter modification scheme in the original reproductive plan. Even though the scheme was based on the reasonable assumption that  $O(t+n)$  gave a better estimate of the optimal genetic operator rate than  $P(t)$ , it had the following obvious faults:

- (i) lack of generality;
- (ii) inability to apportion credit to the operators;
- (iii) consideration of each operator's worth independently of other operators;
- (iv) possibility of a slow and limited search.

In all fairness to Cavicchio, the main aim of his research was not the development of a meta-adaptive plan (parameter modification scheme). Therefore he acknowledged many problems that his meta-plan encountered but he attempted to solve only a few of them. A detailed discussion of the above four faults follows.

(i) To be general the meta-plan must be able to interact favorably with a variety of meta-environments. Cavicchio found however that the meta-plan performed poorly when initial operator rates were low and the adaptive plan used a small population.

Throughout Cavicchio's research his reproductive plan interacted with a deterministic environment which was never modified. Because large steps are undesirable during

latter adaptive stages (with respect to the current environment), the meta-plan tried to force low crossover rates. Once low rates were achieved the meta-plan adjusted these values very slowly (see fault (iv) below). The inability to adjust rapidly to environmental modifications if the operator rates are low is undesirable if the meta-plan is to be considered general.

(ii) Cavicchio was aware of the meta-plan's inability to apportion credit when a new population member was formed using more than one genetic operator. All operators involved were credited equally by the meta-plan. However Cavicchio felt that the crossover operators contributed more than the mutation operators to the development of a new population member. This was not true in all cases. If crossover took place between short pieces of the individuals, or if the parents had similar makeups, then crossover would have effects very similar to mutation. To reward crossover as the main contributor to the formation of a new individual would have been inappropriate. Cavicchio hypothesized that even though the meta-plan was supposed to lower crossover rates they remained high due to crossing-over of similar parents.

(iii) Considering the genetic operators to be independent was another fault. The feedback from the environment shows the ability of combinations of the operators to form new population members. However each operator rate was changed according to its corresponding

operator's performance, regardless of any interactions with other operators. This problem was evident when many good mutations were destroyed by deleterious crossovers (operator performance was recorded only if the individual created using the operator became a new population member). It also occurred when the crossover operator enhanced the mutation operators whose solitary involvement would not have been good enough to have created a new population member (operator involvement was credited equally -- see fault (ii) above).

Cavicchio acknowledged a similar problem in that the inversion rate was a function of both crossover operator rates. Inversion enhances the effectiveness of the crossover operators. Because the inversion rate was the average of the crossover and double crossover rates, it was dependent on the performance of the crossover operators. Eventually crossover took place between similar parental segments. Both crossover rates would have soon risen since the operators would have been similar to mutation. Inversion would have also risen, with a resultant mixing of the gene maps of the new population members. The crossover operators would then have looked like crossover again. The resultant effect, in all likelihood, would have been a lowering of the rates. Cavicchio felt that this method for modifying the inversion rate caused both crossover rates to fluctuate widely.

(iv) The rate of change of genetic operator rates

was highly dependent on the rates themselves. If the operator rates were low or significantly wrong then the appropriate action could have been taken too late (Figure 3.1) or insufficient action could have resulted (Figure 3.2). Thus either the feedback was no longer valid or the space  $\mathcal{A}^1$  was not being searched fast enough.

The rate of change was not the only problem with the search. The search was also limited. The perturbing of rates depended on local data (one set of current rates and one set of observed rates). Hollstien [10] has shown that reproductive plans work much better than plans that rely on local data (because of the properties of a reproductive plan's parallel search). A limited search by the meta-plan is a serious problem if the meta-environment changes. In this instance the meta-environment is constantly changing since the reproductive plan is continually searching new areas of  $\mathcal{A}^0$ , thus creating a need for different search strategies. Thus the meta-plan requires more information than a local (serial) search provides.

### 3.3 A Proposed Reproductive Meta-plan

The genetic operator rate modification scheme proposed here is a reproductive meta-plan. The plan has evolved from Bagley's idea of meta-adaptation [1]. Bagley discussed the use of a meta-adaptive plan that would modify the application rates of those genetic operators used by the adaptive plan.

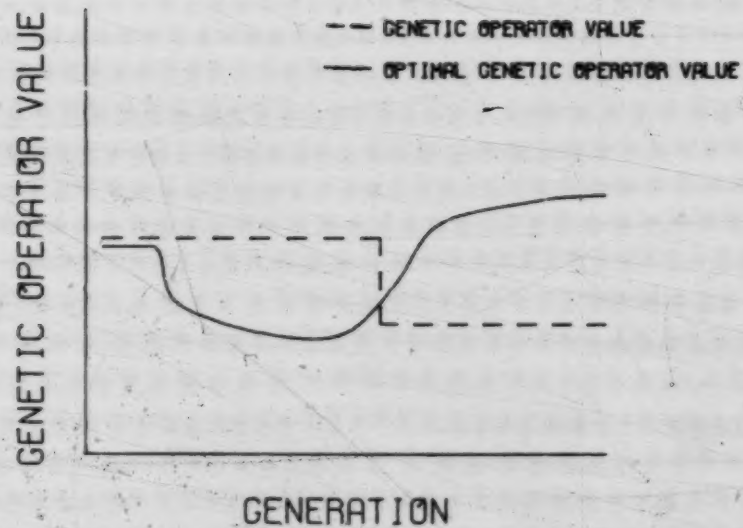


Figure 3.1 Illustration of inappropriate action taken as a result of too many generations per meta-generation.

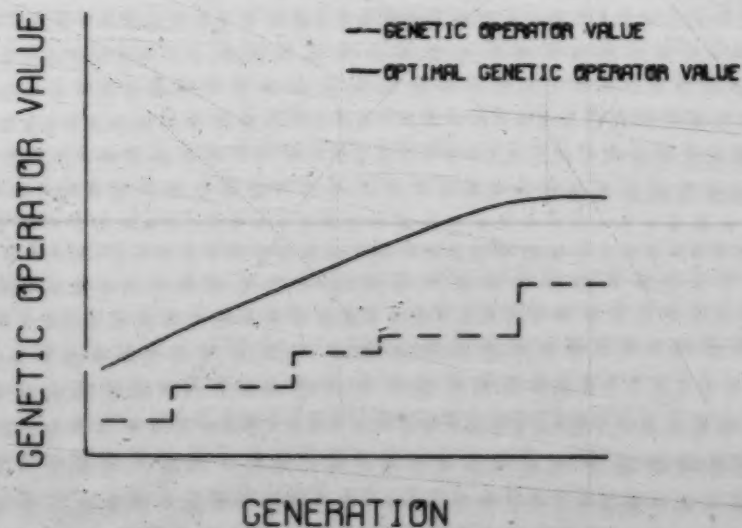


Figure 3.2 Illustration of insufficient action taken as a result of too many generations per meta-generation.

The meta-plan produces strings of genetic operator rates to control the rate of connection (crossover, double crossover, and inversion) and the rate of replacement (mutation). These provisions might suffice. However it would be desirable to study how a non-uniform breakpoint probability distribution affects the rate of finding good structures by the adaptive plan. Thus a system of biasing the breakage points will also be considered.

The proposed meta-plan will be embedded in Cavicchio's reproductive plan [5] for the following reasons. First, an established adaptive plan meta-plan combination can be used as a means of comparison to see if the proposed meta-plan does in fact improve the efficiency and effectiveness of the adaptive plan more than the original meta-plan. Second, Cavicchio's parameter modification scheme has some problems. It would be appropriate to embed the proposed meta-plan in the same adaptive plan to compare these specific problem areas in the event that the desired improvement is not attained. Third, Cavicchio uses an environment that requires a lengthy period of adaptation, which facilitates the study of a meta-plan. The ease with which this environment can be modified slightly also permits the study of the effects of changing environments on both types of meta-plan.

In the reproductive meta-plan experiments that are reported in Chapter 4, the preselection scheme and the reproductive phase of the original  $\pi$  are not changed from

those described in Section 2.5. The difference in the selection scheme is that the sum of the reproduction quotas will now be  $(M - N)/n$ , where  $n$  is the number of meta-structures. As shown in Figure 3.3 each meta-structure (MS) will use the same parents to produce a subset of the offspring. This may not be the best selection method but the main concern is to see how different meta-structures operate. The selection method used will remove some biases that could occur if different parents were used. A utility is assigned to each meta-structure according to how "good"

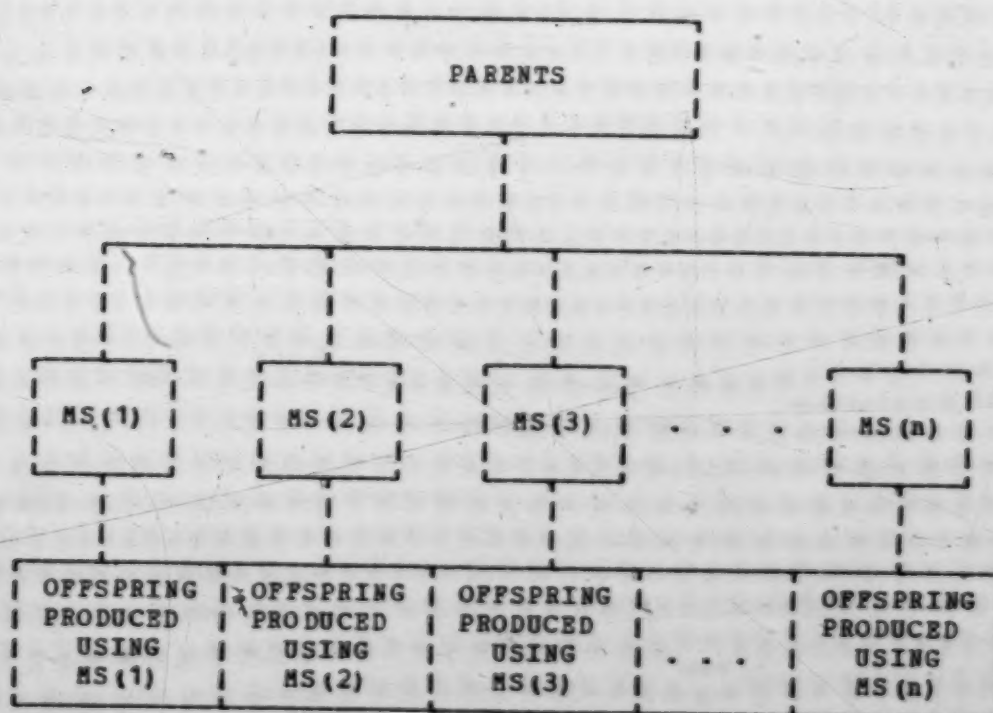


Figure 3.3 Method used by the reproductive meta-plan for producing offspring. MS stands for meta-structure.

its subset of offspring is (this will be described in more detail later).

The proposed meta-plan has been developed within the framework described in Chapter 2. It would be appropriate then to describe the details of the meta-plan in that context.

Each meta-structure,  $A^i \in \mathcal{A}$ , is a combination of elements composed of exactly one rate for each genetic operator used in Cavicchio's reproductive plan together with a group of numbers representing the relative probabilities of breakpoints. These breakpoint values are used by the adaptive plan when it requires a breakpoint for one of its recombination operators or a gene position for one of its mutation operators. The choice of a breakpoint is thus not necessarily uniformly random. Each meta-structure has the representation shown in Figure 3.4.

Since each genetic operator rate is a probability measure the rates are restricted to the interval  $[0.0, 1.0]$ . The minimum rate restriction that Cavicchio imposed on his system is removed. The rate of inversion no longer depends on the rate of the two crossover operator rates. The breakpoint values can range over the interval  $[0, +\infty)$ , although the upper bound is effectively finite. Thus the meta-plan can "disable" a breakpoint. This could enhance the benefits that gene distance provides.

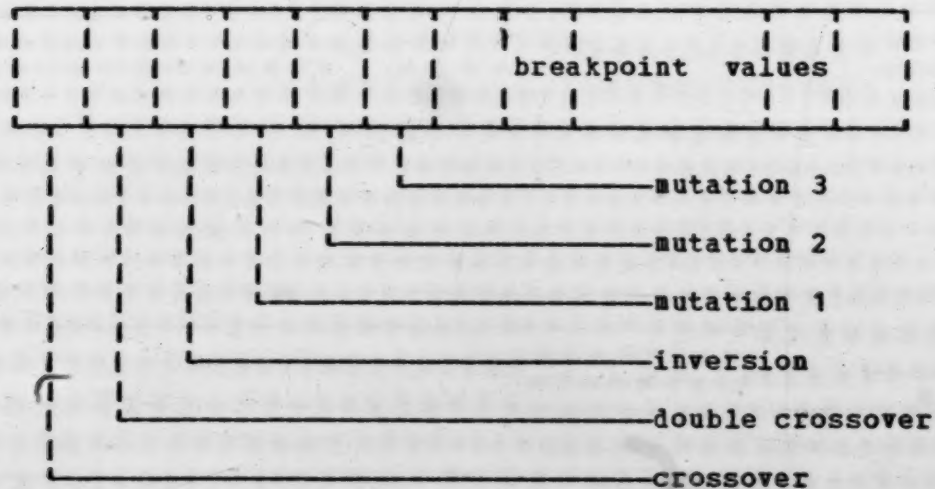


Figure 3.4 Representation of a meta-structure.

In all the experiments described in Chapter 4 the meta-environment,  $E^1$ , was restricted to Cavicchio's reproductive plan (a stochastic environment which is never modified). However the meta-plan could be used with any reproductive plan by modifying the representation of the meta-structures. This could be accomplished by adding positions to and deleting positions from the genetic operator rate region of the meta-structure and changing the number of breakpoint values that is contained in the meta-structure. Thus  $E^1$  is effectively the set of all reproductive plans.

The utility,  $U^1$ , of a meta-structure is the number of possible new population members (in the meta-environment) found in the meta-structure's subset of offspring. The term "possible" is used since this number is obtained

irrespective of the other subsets of offspring. When the new population members are chosen by the adaptive plan all offspring compete; therefore some potential new population members will not survive, not because they are not better than the last generation's but because of the size allowed for the current population,  $N$ . The utility can take on values in the interval  $[0, p]$  where  $p = (M - N)/n$ .

There are many candidates for the utility function. In addition to the one above some are: (1) ranking meta-structures according to the utility of their best individual, (2) ranking meta-structures according to their average utility of their subsets, and (3) ranking meta-structures according to how many actual new population members the meta-structure produces. The proposed utility was chosen because the number of (possible) new population members that are created by each meta-structure (comparisons should be made after the utilities have been calculated) is a reasonable estimator of the ability of each meta-structure to help generate an efficient and effective search in the space  $\mathcal{S}$ . This utility implicitly defines a criterion to estimate the goodness of the individuals produced by the adaptive plan. It seems reasonable to rank the meta-structure higher (in the sense of generating a better search) if it is used to create many "average" new population members rather than a few exceptional ones. (A modification to the function was introduced after some experimentation, as described in

## Section 4.2)

The  $q$  best individuals in the meta-population are kept for reproductive purposes in the next meta-generation. These  $q$  meta-structures and the number of (possible) new population members that each produced comprise the memory,

*M*<sup>1</sup>.

The function  $\tau^1$  generates the next set of meta-structures to interact with the meta-environment. This function tries to generate those meta-structures that will assist the reproductive plan to search *so* more efficiently and more effectively. To do this  $\tau^1$  creates meta-structures in accord with 2 requirements: (1) to provide as wide a range of rates for each genetic operator as possible (allowing quick response to changes in the environment,  $E^0$ ) and (2) to converge on what may be an optimal set of genetic operator rates for the particular search that the reproductive plan requires.

The function  $\tau^1$  uses one type of genetic operator that has not been previously discussed, frame-shift. Frame-shift is an insertion or deletion of a basic building block in an allele. If an operator is viewed as a string of binary digits then a deletion of the rightmost bit and an insertion of a new leftmost bit would be two frame-shifts. If this binary integer is the numerator of a fraction with the denominator all 1's then a 1 inserted in the leftmost bit after a rightmost deletion would be the same as  $(\text{'rate'}+1)/2$  with rounding down and a 0 inserted in a like

manner would be the same as 'rate'/2 with rounding up.

The function  $\tau^1$  forms new genetic operator rates in the meta-structures using multiple crossing-over with frame-shifts (2,3,4,5 below) and without (6 below), and two types of mutation (1,7 below). In the actual implementation one of the following operators is chosen at random to generate each new genetic operator value in each new meta-structure, using  $v(1)$ , the particular operator rate in the best meta-structure, and  $v(2)$ , the rate in the second best meta-structure. The operators are:

- (1)  $v = (v(1)+v(2))/2$ , the average of the rates,
- (2)  $v = (v(1)+1.0)/2$ , half way between  $v(1)$  and 1.0,
- (3)  $v = (v(1)+0.0)/2$ , half way between  $v(1)$  and 0.0,
- (4)  $v = (v(2)+1.0)/2$ , half way between  $v(2)$  and 1.0,
- (5)  $v = (v(2)+0.0)/2$ , half way between  $v(2)$  and 0.0,
- (6)  $v = v(1)$ , the best rate,
- (7)  $v =$  a random rate.

The function  $\tau^1$  also generates new breakpoint values using crossover, inversion and mutation. Crossover and inversion occur with a probability of 0.5. Breaks occur at gene boundaries with equal likelihood, where the genes represent relative probabilities of breakpoints. Each gene mutates with a probability of 0.1. The two mutation operators increment and decrement a breakpoint value by 1. The mutation operator is chosen randomly.

The reproductive meta-plan's operators,  $\Omega^1$ , cannot be identical to the reproductive plan's operators since the

structures and the meta-structures are quite different. The structures are strings of detectors and all the elements of  $\Omega^0$  operate on each locus in a like manner. Because the genes in the meta-structures can be categorized as two distinct types, genetic operator rates and breakpoint values, the elements of  $\Omega^1$  must operate on these two categories in different ways. Thus multiple crossovers and frameshifts (but no inversions) can occur among genetic operator rates, while only a single crossover (along with inversions) is allowed among breakpoint values. Although mutations can occur at all loci, the method used to choose the new allele differs with the meta-locus, as does the rate of mutation.

A criterion function,  $\chi^1$ , will be used to compare Cavicchio's parameter modification scheme and the proposed meta-plan. This function will measure two attributes of the meta-plans: (1) the ability of the meta-plans to search the meta-space in certain situations, and (2) the rate of increase in utility of the population  $\mathcal{B}^0$ . Since  $\tau^0$  generates different numbers of offspring per generation using the different meta-plans,  $\chi^1$  must take into account these differences.

## CHAPTER 4

### EXPERIMENTAL RESULTS

This chapter tabulates the results of computer experiments and proposes some conclusions based on these results. The experiments were run on an IBM 360/67 (using the MTS operating system). The implementation was written in the AlgolW programming language. The experiments were run using the reproductive meta-plan proposed in the previous chapter and with Cavicchio's meta-plan [5]. To make the text less wordy, whenever "meta-plan with (without) preselection" is used in this chapter this is to be interpreted as "the meta-plan applied to the reproductive plan that uses (does not use) the preselection scheme".

Two experiments were run using Cavicchio's meta-plan, one with and one without his preselection scheme. These studies provided some benchmarks for the reproductive meta-plan, and also some indication of the decrease in utility that can be encountered when preselection is used with Cavicchio's meta-plan. Cavicchio decided that preselection had adverse effects on the meta-plan (despite an increase in utility after its addition) because preselection seemed to favor offspring created by crossover more than those created solely by mutation. However in Section 4.1 it will be shown that mutation rates were

decreased by the meta-plan with and without preselection. This favoring of crossover-produced offspring caused crossover rates to increase during the latter stages of adaptation, which Cavicchio regarded as detrimental. However in Section 4.2 a discussion of the plausibility of high crossover rates near the end of adaptation will be given.

Three experiments tested the reproductive meta-plan with preselection. One objective was to determine how a modification of the deterministic  $E^0$  affected both meta-plans. An environmental change occurred when 250 or 260 offspring had been produced. This change consisted of shifting the unknown letters one column right. Another change occurred when 500 offspring had been produced. The modification shifted the unknown letters one column left. Two special experiments were run to test particular items of interest. One tested breakpoint biasing, the other tested Cavicchio's meta-plan with preselection at a "crisis" point.

As described in Section 2.5, the task was to find a set of detectors (binary coded strings of grid coordinates) which would correctly match the unknown letters with the known letters. To review the notation introduced in that section, a population,  $B^0$ , which is  $M/N$  has  $M$  individuals, with the best  $M$  individuals kept as potential parents and  $M - N$  offspring introduced each generation. To describe the reproductive plan embedded in a reproductive meta-plan, an  $M/N(p)$  population will mean that each of the  $M$

meta-structures will be used by  $\tau^0$  to generate  $p$  offspring.

Thus  $p_n = N - N$ .

#### 4.1 Performance of Cavicchio's Meta-plan

Figures 4.1 and 4.2 show the changes in genetic operator rates for a period of 75 generations for the present implementation of Cavicchio's meta-plan with and without preselection respectively. Figure 4.3 shows the utility curves for both experiments. In this graph as in all subsequent presentations of utilities, the utility presented is the average of the best 10 structures in the population.

In replication of Cavicchio's findings the two crossover operator rates in Figures 4.1 and 4.2 show no trends (except that they tend to shift in opposite directions when modified), while the mutation operators converge to rates, in these experiments, between 0.5 and 0.6. (Cavicchio's results show convergence to slightly lower rates; but he ran his experiments for over 200 generations and did not modify the environment.) Figure 4.3 shows that the reproductive plan which does not use preselection outperforms the plan that does, for the first 25 generations and, if the amount of increase in utility is considered, the last 25 generations as well.

Two observations can be made about these results. First, the meta-plan seems to be "unaware" of any changes in the environment. Since the only method for generating new

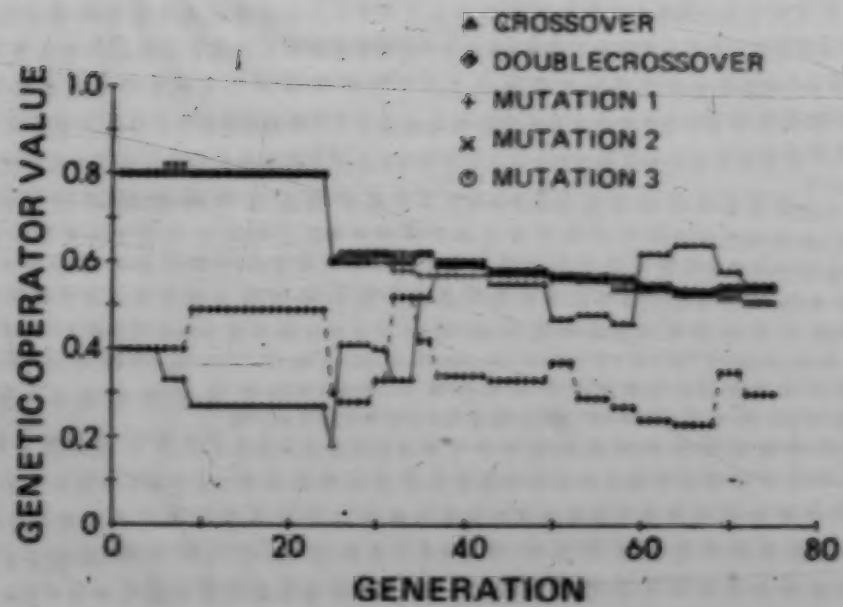


Figure 4.1 Genetic operator rates of Cavicchio's seta-plan, with preselection, for 75 generations.

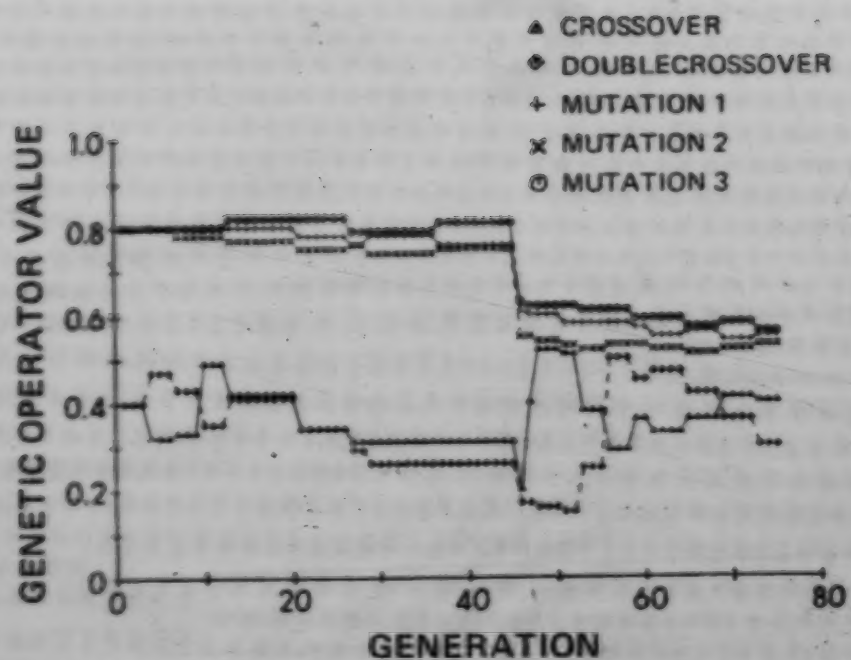


Figure 4.2 Genetic operator rates of Cavicchio's meta-plan, without preselection, for 75 generations.

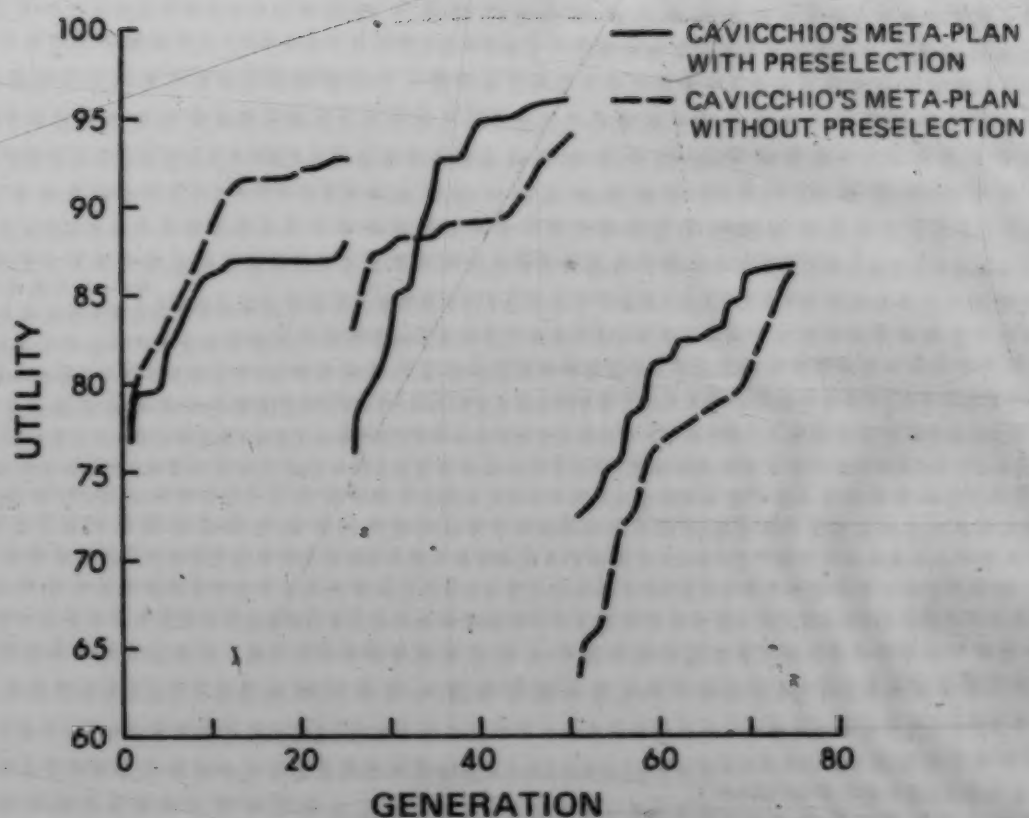


Figure 4.3 Average utility of 10 best structures of Cavicchio's meta-plan with and without preselection.

detectors in this particular reproductive plan is mutation, it would seem reasonable to expect environmental changes to result in increased mutation rates, followed by an increase in both crossover rates. This procedure would inject new genetic material (new n-tuples) into the population and then employ a high rate of recombination to search for the string of n-tuples that performs best. The steady decline in mutation rates thus seems particularly inappropriate.

Secondly the meta-plan is unable to cope adequately with periods of stagnation in the reproductive plan's search. Due to the requirement that a minimum number (N) of new population members must be generated before the genetic operator rates can be modified, these periods of stagnation can cause a loss of efficiency. The method used by the meta-plan to escape this predicament is to halve the required number of new population members. However this action is only taken after the increase in utility over the past 15 generations drops below 3.0. In this experiment that is much too late. At the same time the genetic operator rates are also reduced. (Cavicchio decided that stagnation of the reproductive plan's search meant the operator rates were too high.) In both instances (Figure 4.1 generation 25, and Figure 4.2 generation 46) this appears to be the wrong strategy, since in both cases the meta-plan increases the (single) crossover rate in the following generation.

The other operator rates either increase after a

few generations or decrease immediately. The decrease in mutation rates should not be taken as an indicator since it will be shown in Section 4.4 that this meta-plan seems to be biased towards crossover rate modification and decrease of mutation rates. Fortunately for the reproductive plan, this prescribed decrease in operator rates does not create serious problems, even though the next generation in both cases produces poor results. The number of new population members is already greater than  $N/2$  in both instances, thus another modification of the operator rates will take place in the subsequent generation.

Some indication of the problems mentioned in Section 3.2 has become apparent in these two experiments. The lack of generality is exemplified when the meta-plan reduces the two crossover rates during periods of stagnation, although they should have been increased as the subsequent modification indicates. Also, requiring a minimum number of new population members to be generated before modification of operator rates causes a slow and limited search. Even though the meta-plan halves the number required, this process can happen too late. The second experiment loses 9 generations because the meta-plan is unable to generate new genetic operator rates during this period of time.

#### 4.2 Performance of the Proposed Reproductive Meta-plan

In all experiments described here the meta-population consisted of 5 meta-structures, of which the best 2 remained as meta-parents in the next meta-generation. Table 4.1 shows the best two (parent) meta-structures in

Table 4.1 Initial meta-population and meta-parents for the remaining meta-generations of an experiment using a 60/10(10) population.

| Meta<br>Generation | MS | Cross-<br>over | Double<br>cross-<br>over | Inver-<br>sion | Mutation |      |      |
|--------------------|----|----------------|--------------------------|----------------|----------|------|------|
|                    |    |                |                          |                | 1        | 2    | 3    |
| 1                  | 1  | 0.50           | 0.50                     | 0.50           | 0.80     | 0.80 | 0.80 |
|                    | 2  | 0.40           | 0.40                     | 0.50           | 0.80     | 0.80 | 0.80 |
|                    | 3  | 0.30           | 0.30                     | 0.50           | 0.80     | 0.80 | 0.80 |
|                    | 4  | 0.20           | 0.20                     | 0.50           | 0.80     | 0.80 | 0.80 |
|                    | 5  | 0.10           | 0.10                     | 0.50           | 0.80     | 0.80 | 0.80 |
| 2                  | 1  | 0.40           | 0.40                     | 0.50           | 0.80     | 0.80 | 0.80 |
|                    | 2  | 0.20           | 0.20                     | 0.50           | 0.80     | 0.80 | 0.80 |
| 3                  | 1  | 0.20           | 0.20                     | 0.50           | 0.80     | 0.80 | 0.80 |
|                    | 2  | 0.60           | 0.40                     | 0.75           | 0.90     | 0.80 | 0.90 |
| 4                  | 1  | 0.20           | 0.20                     | 0.50           | 0.80     | 0.80 | 0.80 |
|                    | 2  | 0.60           | 0.40                     | 0.75           | 0.90     | 0.80 | 0.90 |
| 5                  | 1  | 0.60           | 0.88                     | 0.88           | 0.45     | 0.40 | 0.85 |
|                    | 2  | 0.40           | 0.10                     | 0.46           | 0.90     | 0.90 | 0.90 |

each meta-generation (and the entire starting population) for an experiment involving a 60/10(10) population. The 5 crossover-double crossover combinations were initialized to (0.5,0.5), (0.4,0.4), (0.3,0.3), (0.2,0.2), and (0.1,0.1), all inversion rates were set at 0.5, and all 3 mutation operator rates were 0.8 in each meta-structure. This experiment was run for 5 generations and revealed the extreme variations in genetic operator rates that can occur with this plan.

The 60/10(10) population was used as a test case to see if different combinations of genetic operator application rates would in fact produce a significantly different number of (possible) new population members. As the results tabulated in Table 4.2 indicate, in most cases certain combinations produce a significantly larger number of (possible) new population members.

The size of the pilot experiment population was chosen so that minor variations in the results could be seen more easily. However populations of this size are not economical for simulation because of the amount of extra (and in this instance "wasted") computation that is involved. In most cases there are many more (possible) new population members generated than openings for new population members in the next generation. Those that do not become parents in the following generation are lost forever, because the 10 best individuals are replaced almost every generation.

Table 4.2 Number of (possible) new population members per meta-structure and a summary of the average utility of the 10 best structures in a 60/10(10) population.

| Meta<br>Generation | Average<br>utility<br>(10 best<br>population<br>members) | Number of (possible)<br>new population members<br>per meta-structure |   |   |   |   |
|--------------------|----------------------------------------------------------|----------------------------------------------------------------------|---|---|---|---|
|                    |                                                          | 1                                                                    | 2 | 3 | 4 | 5 |
| 1                  | 76.8                                                     | 4                                                                    | 5 | 3 | 5 | 5 |
| 2                  | 82.8                                                     | 1                                                                    | 3 | 1 | 1 | 2 |
| 3                  | 84.5                                                     | 3                                                                    | 3 | 3 | 3 | 3 |
| 4                  | 86.6                                                     | 2                                                                    | 1 | 2 | 3 | 5 |
| 5                  | 87.9                                                     | 1                                                                    | 1 | 0 | 1 | 2 |

Further, since the results will be compared on an offspring basis with the data produced by the experiments discussed in Section 4.1, this meta-plan would be at a great disadvantage. Since the generations would be in a 5 to 1 ratio (Cavicchio's meta-plan : reproductive meta-plan) the reproductive plan using Cavicchio's meta-plan produces 40 out of 50 offspring using successively better sets of parents. Most of the utility gain is attributable to the increase in parental "goodness" not the use of optimal genetic operator rates (the latter only enhances the former). Table 4.2 reveals the relatively poor performance

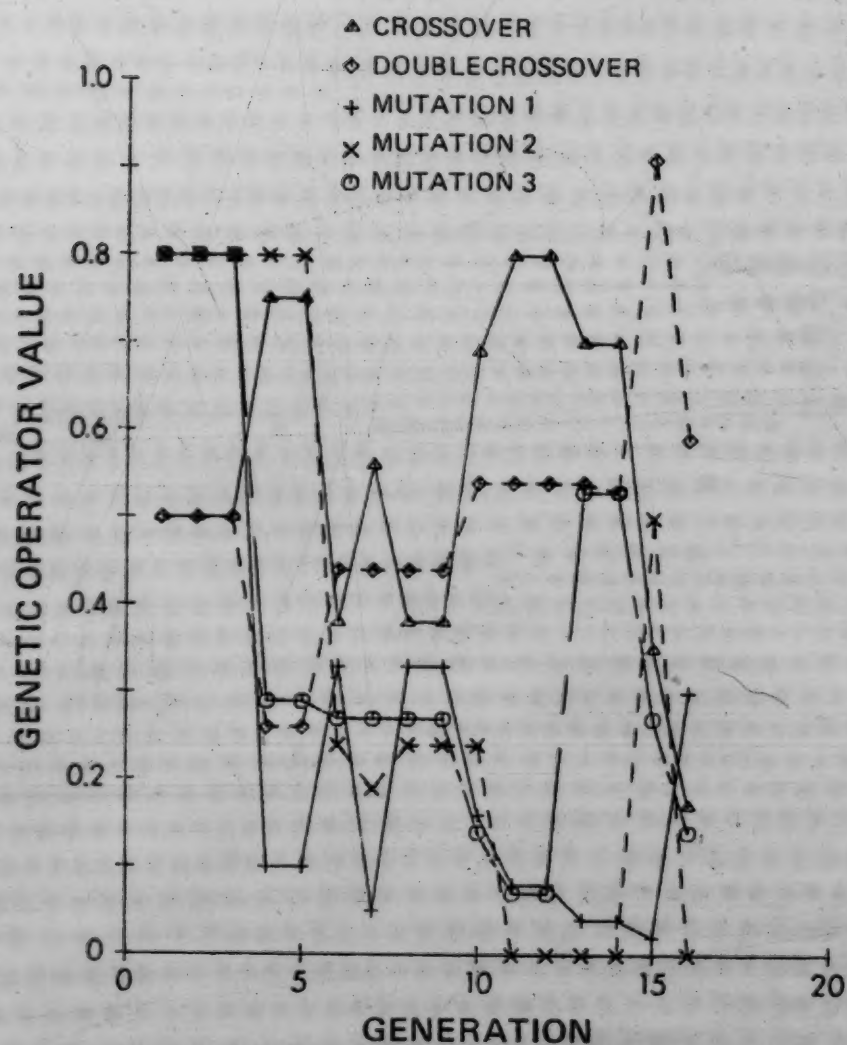


Figure 4.4 Genetic operator rates of original reproductive meta-plan for 16 generations.

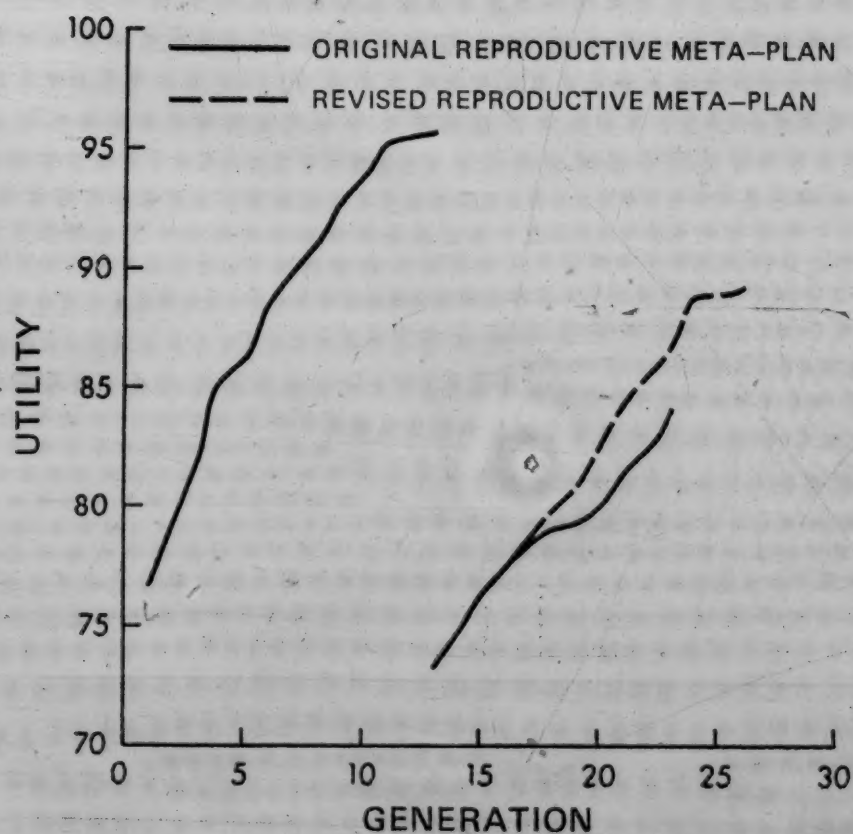


Figure 4.5 Average utility of 10 best structures for original and revised reproductive meta-plans for 27 generations.

of this population size when compared with Cavicchio's meta-plan without preselection at generation 25 (Figure 4.3 shows an average utility of 92.6).

For reasons of economy and fairer comparison it was decided to use smaller populations (30/10(4)) in the subsequent experiments. Figure 4.4 shows some surprising results concerning the best meta-structure. This and all subsequent presentations of genetic operator rates will be concerned with the best meta-structure only. As the population of structures approaches a matching score of 100 (see the utility rates for this experiment in Figure 4.5) the crossover rates stay very high and the mutation rates drop in some cases to 0. This seemingly inappropriate action could have a very good explanation. The population,  $B^0$ , at this point has been partitioned into four distinct groups. Each partition includes a number of individuals with minor variations in genetic content. What might be happening is that crossover and double crossover have become equivalent to non-random mutation. The pieces exchanged between two structures are so similar that the effect is like mutation, except that it is not random and therefore has a higher likelihood of improving the old individuals. Apparently, the "problem" Cavicchio associated with late high crossover rates may in some cases be a disguised advantage.

To see what trends might develop with the genetic operator rates (since an environmental change is scheduled

at this point) this experiment is continued for three more generations (generations 14, 15, and 16 in Figures 4.4 and 4.5). No trends appear in the crossover rates and there is a slight increase in the mutation operator rates, especially those that cause the smallest changes (mutation 1 and mutation 2). These changes are probably due to the fact that crossover and double crossover cannot supply new genetic material but do not harm the formation of new individuals since their effects are negligible at this point; and the "tuning in" property of mutation 1 and mutation 2 is enhanced.

An environmental change occurs at generation 14 (it occurs here since it is desirable to have the environmental changes in all experiments coincide (approximately) on an offspring scale rather than a generation basis). The results shown in Figure 4.5 are evidence of a comparatively worse performance by the reproductive meta-plan than by Cavicchio's meta-plan with preselection. One explanation is that at this point the meta-plans are working with two very different meta-environments. With Cavicchio's meta-plan *B<sup>0</sup>* retained much more of its initial variance than with the reproductive meta-plan. (In the latter case, the reproductive plan was working with only 82 different alleles, out of a possible 430.)

An oversight during the development of the reproductive meta-plan is apparent in some of the results.

The difficulty is with the method of choosing the best meta-structures when there are ties involving the number of (possible) new population members. The method used for breaking these ties was to take the first two from the list of meta-structures that were tied. This method results in keeping the present meta-parents if none of the meta-offspring were better. However, as the instance presented in Table 4.3 shows, a meta-structure involved in a tie is discarded even though it is used to generate better offspring.

Table 4.3 Meta-generation 17 showing the selection method for the original and revised reproductive meta-plans.

| Rank of<br>$A^1$ in<br>$B^1(t)$ | Number of<br>possible<br>new<br>population<br>members<br>produced<br>using $A^1$ | Utility of<br>possible<br>new<br>population<br>member(s) | Rank of<br>$A^1$ in<br>$B^1(t+1)$<br>(original $\tau^1$ ) | (revised $\tau^1$ ) |
|---------------------------------|----------------------------------------------------------------------------------|----------------------------------------------------------|-----------------------------------------------------------|---------------------|
| 1                               | 2                                                                                | 77.7 77.3                                                | 1                                                         | 2                   |
| 2                               | 1                                                                                | 77.3                                                     | -                                                         | -                   |
| 3                               | 2                                                                                | 77.3 77.3                                                | 2                                                         | -                   |
| 4                               | 0                                                                                | -                                                        | -                                                         | -                   |
| 5                               | 2                                                                                | 78.1 77.8                                                | -                                                         | 1                   |

The following revision was made to the reproductive meta-plan. The meta-structure chosen in case of a tie was the one which produced the structure with the highest utility. Table 4.3 shows the first instance (meta-generation 17) of a different choice of meta-structures. It is interesting to note that in the original reproductive meta-plan the meta-structure that is labelled "1" was never replaced for the remainder of the experiment even though it was involved in a tie in half of the remaining meta-generations. In the revised meta-plan however this meta-generation was the beginning of a trend towards higher mutation rates. It appears that as the reproductive plan finds it harder to produce significantly improved structures the meta-plan finds it harder to replace the meta-structure that is part of the cause of the poor performance by the reproductive plan.

The resulting gains in utility obtained by this revised reproductive meta-plan are evident in the results presented in Figure 4.5. A significant improvement is evidenced by the increase in utility,  $\mu^0$ , after the first few generations following the environmental change. (Meta-generation 17, which occurs between generations 16 and 17, was the first instance of a different choice of a meta-structure involved in a tie. Since the utility that is plotted is that of the population before reproduction the increase doesn't appear until generation 18.) Three things should be particularly noticed. First the increase in

utility is not as large as for the reproductive plan using Cavicchio's meta-plan with preselection, possibly because the reproductive plan using the reproductive meta-plan starts with the disadvantageous low population variance. Second the revised reproductive meta-plan makes a somewhat greater gain in utility than does Cavicchio's meta-plan without preselection. Third the mutation rates, shown in Figure 4.6, increase as expected (after a slight time lag), possibly showing the meta-plan's "awareness" of the problem at hand. In order to improve, the meta-plan must allow the reproductive plan to generate new genetic material. The significant improvements in  $\mu^0$  do not appear until the mutation rates have increased. This is the "crisis" point which will be discussed further in Section 4.3.

Since the other four meta-structures in the meta-population have not been mentioned, it would be appropriate to discuss what trends occurred with them. Since meta-structures 3, 4, and 5 are new for every meta-generation the individual genetic operator rates tended to be random except when the particular rates in the parent meta-structures (1 and 2) are similar. In this case the rate is biased slightly towards the rate in the parents, since one of the meta-operators averages the two parental rates. And the rate is even more biased when the rates in both meta-parents are close to 0 or 1, since four meta-operators force the rates closer to 0 or 1.

The most important of these four meta-structures

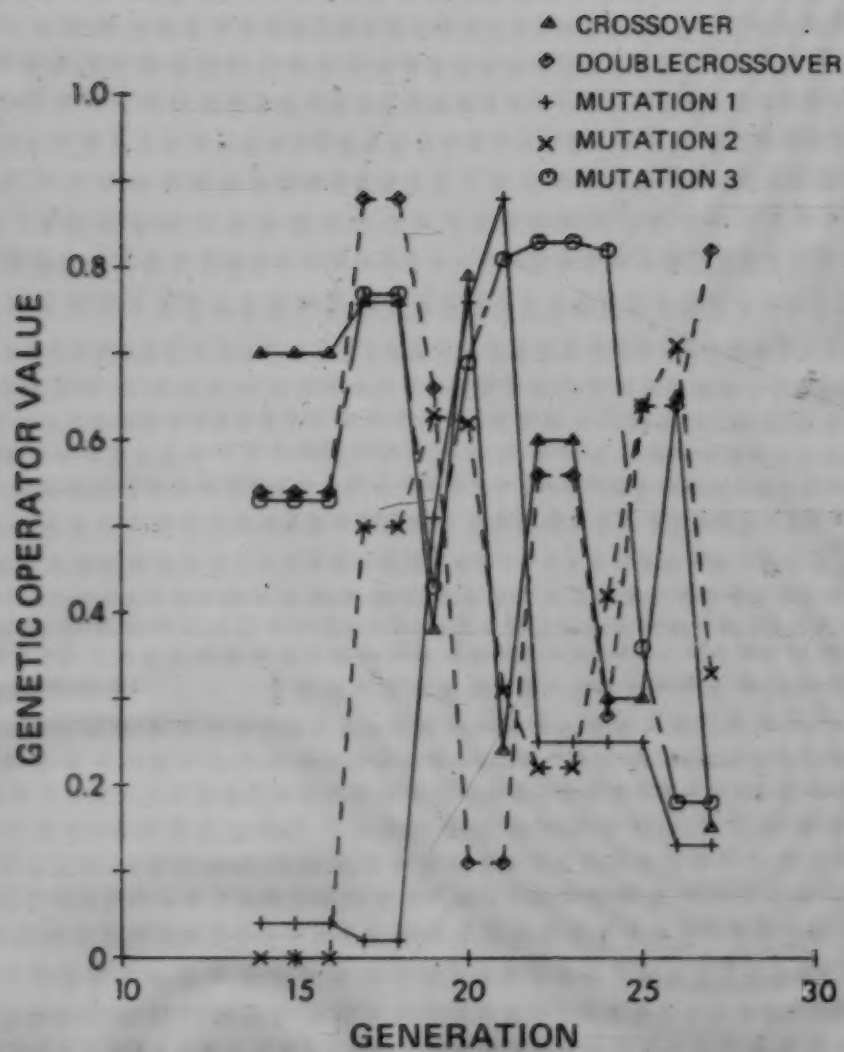


Figure 4.6 Genetic operator rates of revised reproductive meta-plan for 14 generations.

is meta-structure 2 since it is one of the meta-parents. The trends that occurred with this meta-structure generally coincided with meta-structure 1, usually with some time lag. Two explanations for this similarity are: (1) meta-structure 1 occasionally was replaced but was good enough to be retained as the second parent, and (2) all of the meta-operators that are functions of meta-structure 1 reflect the change that occurs in this meta-structure. Therefore as meta-structure 1 starts reproducing, some of the meta-offspring tend to mimic the trend that occurred when this meta-structure became a parent. If one of these offspring becomes meta-structure 2 the trend will appear and will be consistent with the time-lag phenomena.

The above discussion emphasizes the "drift" of the meta-population in the same direction as the best meta-structure. However, the equally important inherent ability of the meta-plan to explore different areas of the space of meta-structures coexists with the ability to follow the best meta-structure's trends, because some of the meta-operators force the search in different directions.

#### 4.3 Results of an Experiment Involving a "Crisis" Point

Throughout Cavicchio's discussion of the merits of modifiable genetic operator rates [5] he frequently observed that "reasonable" operator rate settings were appropriate at most times. In the following discussion the exceptions will be called "crisis" points.

The "crisis" point of main concern to Cavicchio occurred when the reproductive plan neared its potential capability, that is when it had reached a point where the increase in utility had decreased due to the plan's inability to find significantly better structures. He felt that unduly high crossover rates were destroying potentially good mutations. Therefore the main purpose of his meta-plan was to increase the reproductive plan's potential capabilities by decreasing both crossover rates. He successfully lengthened the plan's effectiveness even though the meta-plan was unsuccessful at decreasing the crossover rates.

Apart from this failure to decrease the crossover rates, Cavicchio's meta-plan failed to meet the criteria of generality. One crisis point not considered by Cavicchio (because he never included the situation in his experiments) was a change to the environment,  $E^0$ . In the present experiments run using Cavicchio's meta-plan all instances of environmental changes occurred when the genetic operator rates were still reasonably high. The worst case that could face the meta-plan at an environmental change is if the genetic operator rates were low.

An experiment was therefore devised to test the effects of this crisis point on Cavicchio's meta-plan. The crossover and inversion rates were set at the 0.1 minimum and the mutation rates were set at the 0.2 minimum. The population of structures that the reproductive plan used was

the population that existed at generation 50 in the second experiment above.

Figure 4.7 shows that Cavicchio's meta-plan is incapable of coping with this "crisis" point. It appears unable to raise the mutation rates, a step which is quite important after an environmental change. Figure 4.8 shows the inability of this meta-plan to increase the utility of the population as quickly as the same meta-plan that starts with more "reasonable" genetic operator rates.

During the experiment with the reproductive meta-plan, (see Section 4.2) the revised reproductive meta-plan proved itself capable of adjusting to a similar "crisis" point (see Figure 4.5 and Figure 4.6). In this case the initial rates for mutations 1, 2, and 3 were approximately 0, 0, and .5 respectively. Not only were these rates low but the population had also lost nearly all of its variance. Yet the meta-plan was able to raise the mutation rates enough to allow the reproductive plan to increase the population variance and to adapt significantly, given the initial problem of low variance.

#### 4.4 Results of an Experiment Done on Breakpoints

Most of the work surveyed in Chapter 2 faced the problem of trying to increase initial performance while keeping the allele loss rate at a minimum and still allowing effective adaptation throughout the observed period. Most of the techniques involved finding appropriate genetic

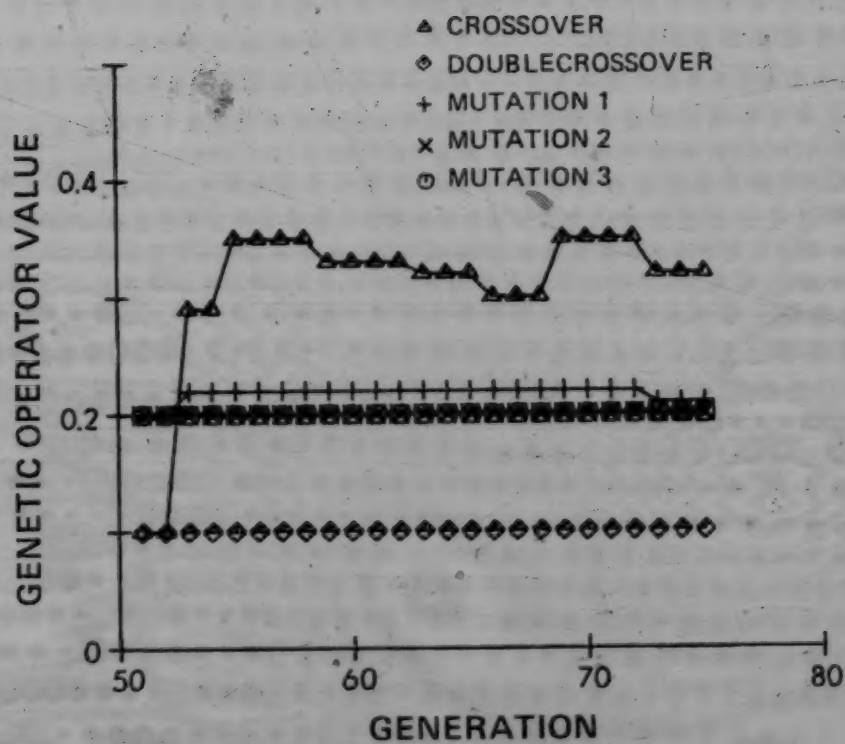


Figure 4.7 Genetic operator rates of Cavicchio's meta-plan, with preselection, at a "crisis" point.

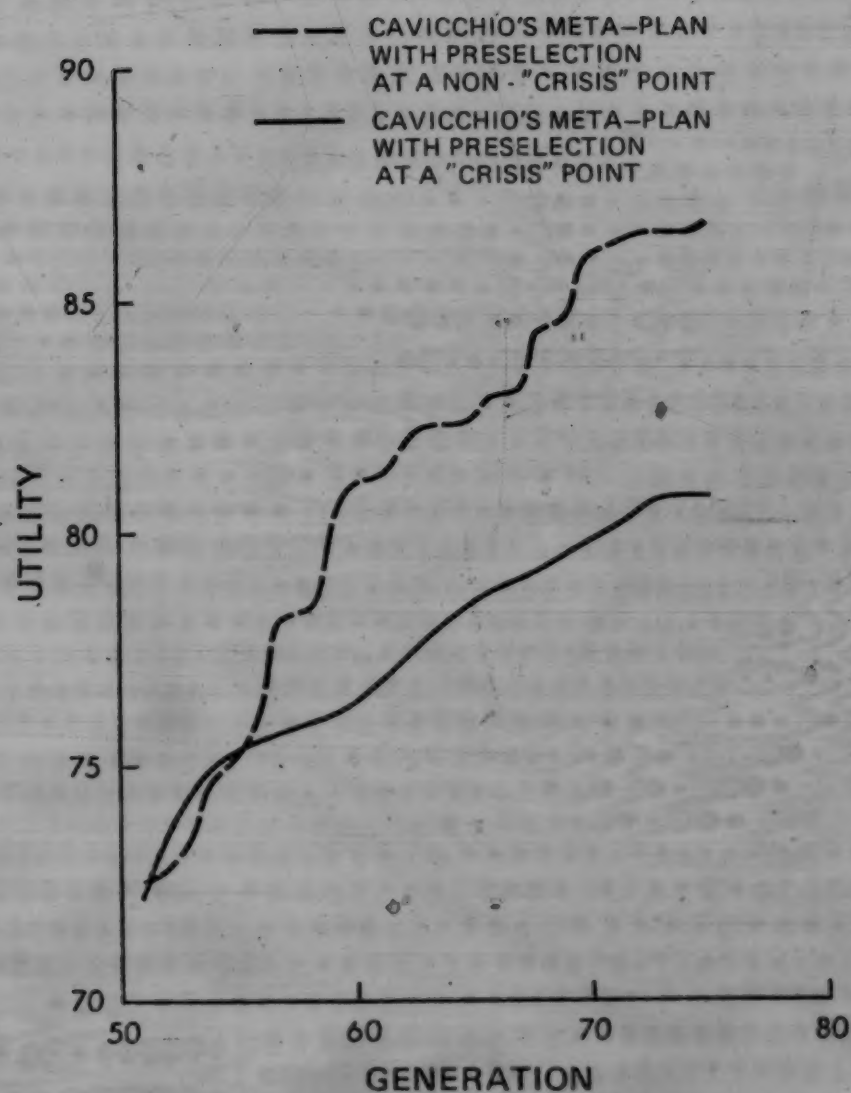


Figure 4.8 Comparison of average utility of Cavicchio's meta-plan, with preselection, at a "crisis" point and a non-"crisis" point.

operator rates that improved the performance to an acceptable level. However some tradeoffs still occurred, such as that between high mutation rates and poor late performance. Hopefully, using a reproductive meta-plan to automatically adjust the rates throughout the adaptive period will reduce the number of tradeoffs.

A less conservative search can be obtained without adjusting the genetic operator rates if the choice of breakpoints for the crossover operators is biased. The most dramatic departures in offspring structures occur when the parents exchange half of their genetic material. Biasing the choice of breakpoints towards the middle of the chromosome should therefore help to increase the search breadth of the plan in the early stages of adaptation.

Since no distinct changes in the relative frequency distributions for breakpoints were noticed during the limited number of generations that were run using the proposed reproductive meta-plan, an experiment was designed to show the significance of different breakpoints on the chromosome structure. The lack of a clear trend could have been caused by an insufficient number of meta-generations or by the inability of the reproductive meta-plan to change the relative breakpoint frequencies quickly enough to have been important to the overall strategy.

Table 4.4 shows the results of these experiments. Three different combinations of crossover and double crossover rates were used. Mutation and inversion rates

were set to 0 since they are irrelevant in these experiments. The relative breakpoint frequencies were biased in two different ways: (1) all breaks occurring in the middle 15 gene boundaries, and (2) all breaks occurring in the first and last 10 gene boundaries. Actually these broad categories are approximate since chromosomes varied from 39 genes to 47 genes in length,

Table 4.4 Summary of breakpoint biasing experiment using two 70/10(10) populations (one for biased meta-structures, and one for unbiased ones).

| Cross-over | Double cross-over | Bias                | Number of possible new population members | Average utility of possible new population members |
|------------|-------------------|---------------------|-------------------------------------------|----------------------------------------------------|
| 0.5        | 0.5               | middle              | 8                                         | 79.7                                               |
|            |                   | end                 | 5                                         | 78.0                                               |
|            |                   | unbiased            | (8,7) <sup>1</sup>                        | 79.2                                               |
| 1.0        | 0.0               | middle <sup>2</sup> | 7                                         | 79.7                                               |
|            |                   | end                 | 7                                         | 78.8                                               |
|            |                   | unbiased            | (6,5) <sup>1</sup>                        | 78.7                                               |
| 0.0        | 1.0               | middle              | 6                                         | 80.7                                               |
|            |                   | end                 | 5                                         | 77.2                                               |
|            |                   | unbiased            | (7,5) <sup>1</sup>                        | 79.6                                               |

<sup>1</sup>Two samples

<sup>2</sup>Because of the better average utility this bias is considered better.

whereas the total length of the breakage point frequencies in the meta-structure is set at 50. The chromosomes' gene boundaries are mapped 1-to-1 against the breakpoint frequencies in the meta-structure. Thus the chromosome that has 39 genes would never map against the 10 places at the end of the meta-structure.

Despite this variation in the meaning of "ends", the meta-structures that have breakpoints totally biased near the "middle" never do worse and frequently do slightly better than those with breakage biased near the "ends". Also the unbiased meta-structures never outperform the "middle" meta-structures and do poorer than the "end" ones only once. This is in keeping with the hypothesis that the largest reproductive changes (most extensive search of the space A) occur when the largest amount of "genetic material" is exchanged between two chromosomes during crossover (which occurs when half of each chromosome is exchanged).

These data are insufficient in both sample size and quantitative differences to warrant any conclusions; however, it appears that there is some evidence to support a more detailed study of the subject of biased breakpoints.

#### 4.5 General Discussion

Figure 4.9 summarizes some of the results obtained from the experiments described in the previous sections. The x-axis has been changed from generations to number of offspring since  $\tau^0$  with the reproductive meta-plan produces

20 offspring/generation whereas  $\tau^0$  with Cavicchio's meta-plan produces only 10. The utility curves for Cavicchio's meta-plan are the same as those in Figure 4.3. The curve for the reproductive meta-plan is the same as Figure 4.5 during phase 1 (between 0 and 260 offspring). Phase 2 (between 260 and 520) shows the performance of the revised reproductive meta-plan (also from Figure 4.5). The final phase (the last 260 offspring) represents the results of an experiment which used the revised reproductive meta-plan. It was started with the same meta-environment and genetic operator rates as Cavicchio's meta-plan with preselection used at generation 51.

As previously discussed the reproductive meta-plan outperforms the nonreproductive meta-plan during the first phase. In the second phase it appears that the reproductive meta-plan outperforms the nonreproductive meta-plan without preselection despite the initial difficulty of a population with little variance. However in the third phase the reproductive plan does very poorly. The reproductive plan seems to suffer more from the initial low variance and thus does not achieve the sorts of gains obtained by Cavicchio's meta-plan in this phase. This problem could be solved by changing the selection scheme used, as discussed in the next chapter.

It appears that there is an overall improvement in searching the space of meta-structures when using a reproductive meta-plan. Initial constraints placed on the

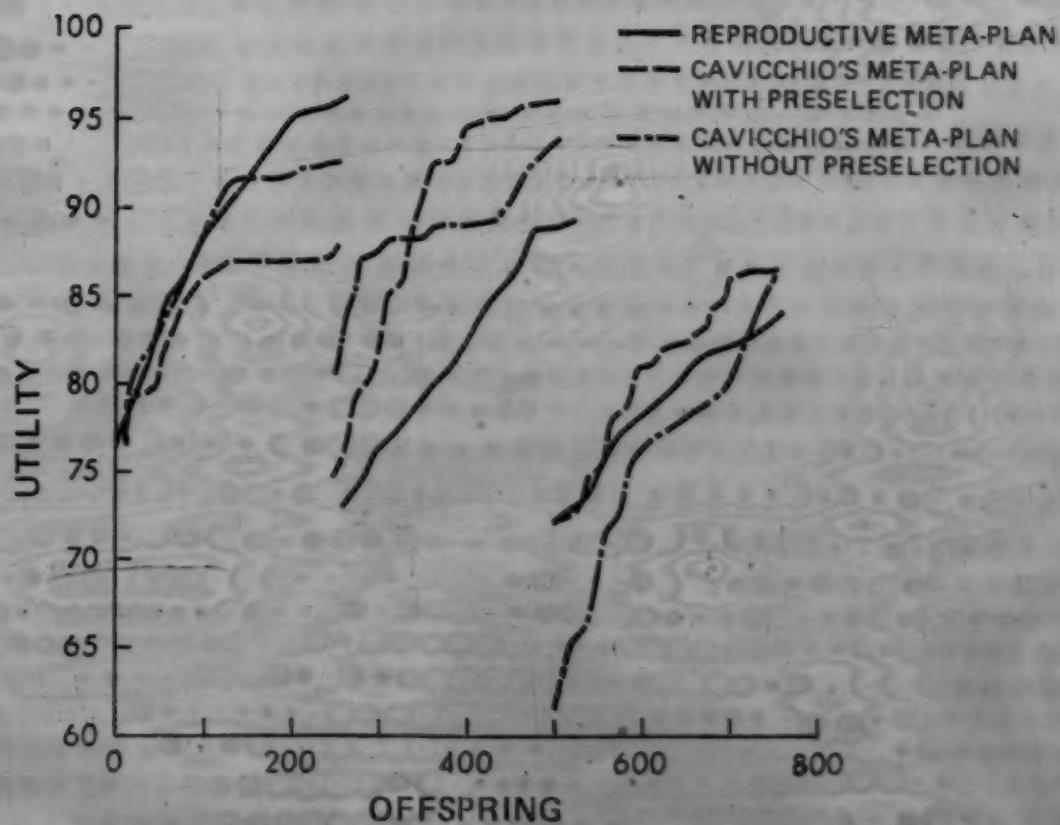


Figure 4.9 Summary of average utility of the 10 best structures for Cavicchio's meta-plan with and without preselection and the original and revised reproductive meta-plans.

nonreproductive meta-plan that cause poorer results show the importance of generality of reproductive plans (see Chapter 2) whenever an adaptive system is implemented. For example, the reliance of the nonreproductive meta-plan on the number of new population members was one cause of a limited search by the meta-plan. This is apparent when the occasional "stair step" effect is seen in the utility curves for Cavicchio's meta-plan (Figure 4.9). And a lack of generality is noticed when the nonreproductive meta-plan decreases operator values whenever there is a stagnation of the adaptive plan.

The first differences plotted in Figure 4.10(b) and Figure 4.10(c) better reveal this lack of generality. In most cases the peaks in the first difference curves for Cavicchio's meta-plan coincide with a change in the genetic operator rates. This is followed by an immediate drop in utility gain. Note that a large increase in utility is more often the cause rather than the effect of a modification of the operator rates. Thus the resultant change which is based on observed frequencies of operator usage to form new population members is suspect with respect to its validity.

Even though there are peaks in the first differences of the reproductive meta-plan (Figure 4.10(a)), these peaks usually are not followed by a dramatic decrease which is the more common occurrence in Figures 4.10(b) and (c). (The peaks immediately following the environmental change are normal because the greatest utility increase

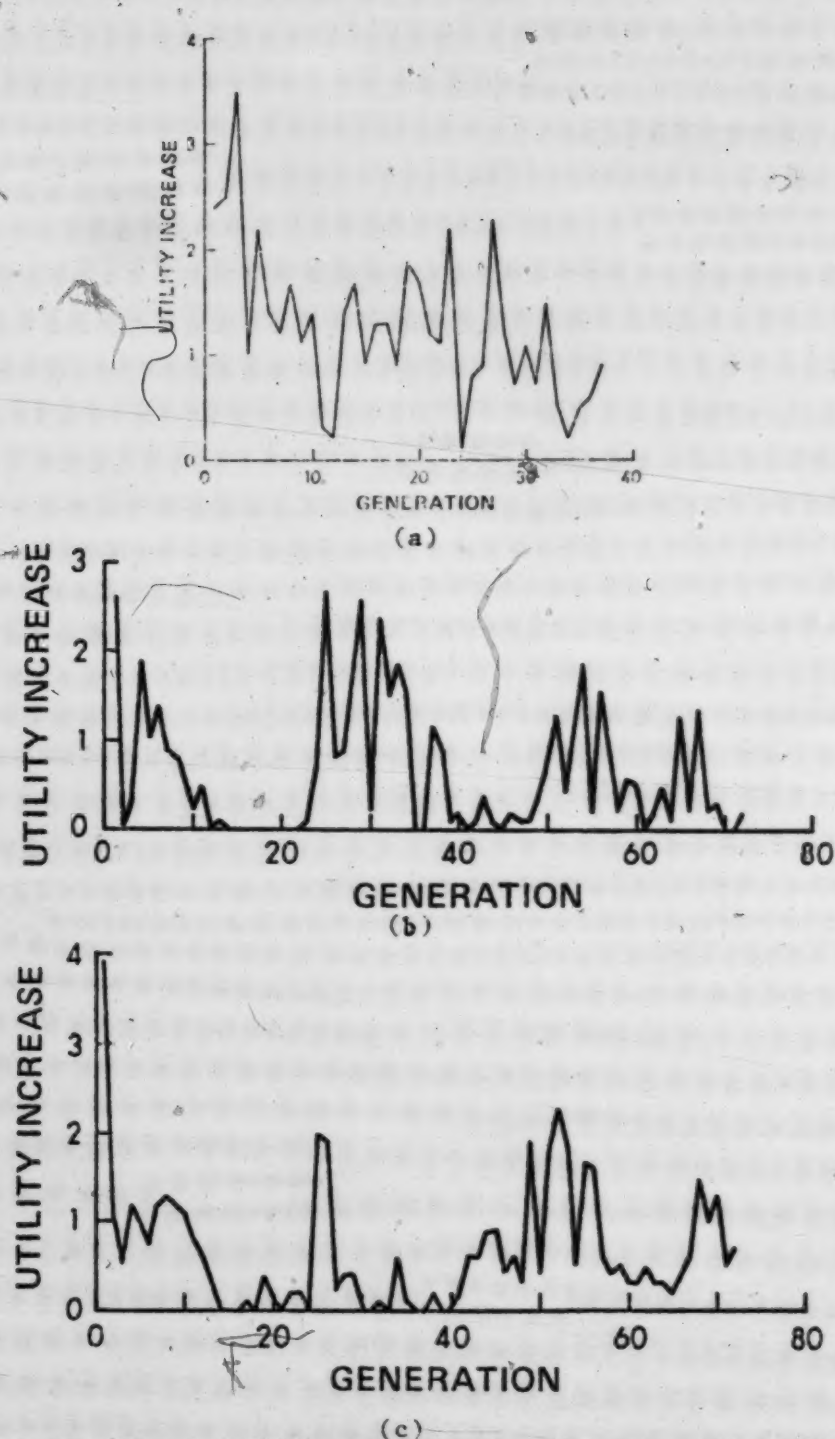


Figure 4.10 First differences of the average utilities summarized in Figure 4.9.

- (a) Reproductive meta-plan.
- (b) Cavicchio's meta-plan with preselection.
- (c) Cavicchio's meta-plan without preselection.

usually occurs when the utility is low.) Figure 4.10(a) does reveal a steady increase in utility caused by the use of nearly optimal genetic operator rates every generation with the reproductive meta-plan, instead of the erratic increases evident in Figures 4.10(b) and (c) caused by genetic operator rate modification every three generations (approximately) with Cavicchio's meta-plan.

Another improvement associated with the reproductive meta-plan is that it seems to be "aware" of environmental changes. The "crisis" experiment shows that the nonreproductive meta-plan is almost totally unable to increase the mutation rates. The reproductive meta-plan on the other hand is able to modify sufficiently the genetic operator rates in such a situation so that the genetic variance of  $B^0$  increases significantly in the second phase.

It should also be noted that mutation rates produced by the reproductive meta-plan fluctuated much more than those produced by Cavicchio's meta-plan. This change from, and the occasionally immediate return to, a certain mutation rate could be due to the method of choosing the number of mutations (see Section 2.5). Large stochastic side effects are intrinsic to this method since two samples of two different random variables govern the total number of mutations on the chromosome. One way to prevent this fluctuation might be to use the method of mutation described in Section 2.2.3, which interprets the mutation value as the

probability of mutating each gene of the chromosome. The number of samples is equal to the number of loci on the chromosome. The stochastic side effects should decrease with this method.

Another explanation of the mutation rate fluctuations is that the reproductive meta-plan is less sensitive to combinations of operator rates than Cavicchio's plan. Thus less important operators can have their rates fluctuate with no apparent loss in the effectiveness of the combination of rates used by  $\tau^0$ .

## CHAPTER 5

### CONCLUSIONS AND FUTURE STUDIES

The primary objective of this investigation has been to develop a meta-adaptive plan that controls genetic operator rates and relative breakpoint probabilities in a first level reproductive plan.

To review the results discussed in Chapter 4, the reproductive meta-plan appears to be superior at "crisis" points. In the non-"crisis" situations studied there is mixed success. In phase 1 (see Section 4.5) the reproductive meta-plan is superior to both instances of Cavicchio's meta-plan. In phase 2, if utility gain is considered, the reproductive meta-plan's performance is between that of Cavicchio's meta-plans with and without preselection. In phase 3 the reproductive meta-plan does slightly worse than Cavicchio's meta-plan with preselection and much worse (considering utility gain) than Cavicchio's meta-plan without preselection. However some of the poor performance can be attributed to the decrease in population variance in the meta-environment. This item will be discussed later.

With emphasis on the superiority of the reproductive meta-plan at "crisis" points, which are the situations most detrimental to the first level reproductive

plan, and on the lack of significant overall superiority of either meta-plan in non-"crisis" situations, it would be fair to say that the reproductive meta-plan is an improvement over Cavicchio's meta-plan. Despite this improvement there remain some significant problems with the results obtained and with the reproductive meta-plan itself. Suggestions for solving each of these problems could direct further research into some interesting areas.

As a result of the extremely high cost of simulation, all of the experiments discussed in Chapter 4 were run only once. Each experiment should ideally have been replicated a number of times with different initial populations and different parameter settings, so that more reliable performance estimates could have been obtained. The first suggestion, then, is that the use of a dedicated computer (which was not available at the outset of this study) should be considered. A perfect candidate would be the Department of Computing Science's microprogrammable QM-1 (Nanodata Corporation) since it could be configured to appear like a population of structures and to have high level instructions like "REPRODUCE", and "MUTATE". Fruitful continuation of the present research could thus involve development of specialized, microcoded software.

A second suggestion is that, under the reproductive meta-plan, the selection method used by  $r_0$  should allow the  $M - N$  offspring to be allocated to all  $N$  parents. The present selection method allocates  $p$  offspring

to some subset of the  $N$  parents, a subset which is used  $n$  times, once for each meta-structure. As mentioned in Section 4.5, this selection method is probably what caused the considerable loss in population variance. One of the main purposes of Cavicchio's selection scheme was to limit the number of offspring of any parent to two per generation. But in the case of the reproductive meta-plan a parent in  $B^0$  can contribute to  $2n$  offspring. Also, if each parent has a reproduction quota of one, then in Cavicchio's reproductive plan, each parent contributes to one offspring. Under the reproductive meta-plan each of the  $p$  parents contributes to  $n$  offspring. The reason for using the same set of parents for each meta-structure was to keep any biases attributable to parental "goodness" from entering into the choice of the next meta-generation's meta-parents. If meta-structures were associated with different parents, the development of a different feedback function to counteract the biasing of the meta-structural utility attributable to parental "goodness" (in  $B^0$ ) would be required. One candidate for  $\mu^1$  would be an averaged relative utility gain.

It should be noted however that the rapid convergence to good structures associated with this poor selection method, even though it caused bad results in phase 3, probably did not cause the significant improvement in phase 1. In all cases in phase 1 where the parents were the two best structures in  $B^0$  the results of poor selection

were an average gain in utility, the incorporation of many similar individuals into  $B^0$ , and a subsequent drop in utility gain from its previous values.

The third change could be in the techniques to counteract these biases when selecting meta-parents. If the parents are allowed to be different for each meta-structure then the validity of the feedback must be preserved. One way to do this is to change the correspondence of generations to meta-generations from 1-to-1 to 2-to-1 or more, thus increasing the likelihood of the meta-structures being compared on a less biased choice of parents. However once the meta-generation is lengthened the meta-plan becomes less sensitive to meta-environmental changes. A loss of sensitivity could be a serious problem since the meta-environment changes every generation (the population,  $B^0$ , is changing). Greater changes occur when the environment,  $E^0$ , changes. The slowness of Cavicchio's meta-plan's search could be attributable to a lack of sensitivity (ratio of generations to meta-generations).

Another improvement might be to fit the number of breakpoint values to the length of the chromosome by removing a sufficient number of loci from the meta-structure. Thus if the breakpoint values bias breakage in the middle of chromosomes, a short chromosome will not be biased near the end (see Section 4.4). This would not be a problem if the chromosomes were all the same length. Also the same frequency distribution should not be used for the

mutation operators and the recombination operators. Intuitively it seems inconsistent with the criteria of a general reproductive plan to make the mutation frequency of a gene the same as the breakpoint frequency next to it. A better control would be to incorporate in the structure an analogue of mutagenic genes [1], or to incorporate a breakpoint value type of component into the meta-structure to specify the relative gene mutation probabilities. Specific studies of methods to effect faster modifications of the relative probabilities could be undertaken.

Other facets of the reproductive meta-plan that do not arise from the above problems could also be studied. For example, biasing the number of offspring produced using a meta-structure according to whether or not it is a meta-parent could result in some improvement. Cavicchio's meta-plan uses the "best" set of genetic operator values to produce all of the offspring. On the other hand the reproductive meta-plan uses the superior meta-structures to produce fewer than half of the offspring and the other meta-structures (which could be inferior) to produce the rest. Another item to investigate could be the development of different meta-operators. Perhaps more of the meta-parents' "genes" could be left intact and more "natural" meta-operators could be used to produce the meta-offspring.

This concludes the discussion of the reproductive meta-plan developed and studied in Chapters 3 and 4.

However this is only one facet of Holland's theory of adaptation. The following items give some indication of the areas of the theory that are still relatively unexplored.

First, adaptation of the utility function to enrich the environment is an important concept. A rich environment is one that has many levels of "problems", from the most general level to specific instances of the "problem". One method of enriching an environment is the generation of auxiliary utility functions which check for solutions of subgoals (specific instances of the "problem"). The structures that solve these subgoals can then be used to generate more general structures to solve higher level subgoals. Second, the representation of structures should be adaptive so that the representation is not a constraint on the plan's adaptive ability. Third, the ability to generate new operators automatically would be desirable. For example, in the pattern recognition task a mutation operator that shifts an n-tuple to a coordinate-wise adjacent n-tuple might have desired effects at certain points in the adaptive process. And fourth, the use of other evolutionary mechanisms such as dominance could fruitfully be studied. Dominance combined with a dominance change operator is a very effective method of storing schemata as recessives or releasing schemata as dominants in one operation.

U

## REFERENCES

- [1] Bagley, J.D. The Behaviour of Adaptive Systems Which Employ Genetic and Correlation Algorithms. Ph.D. Dissertation, University of Michigan, 1967.
- [2] Bellman, R. Adaptive Control Processes: A Guided Tour. Princeton University Press, 1961.
- [3] Bethke, A.D., B.P. Zeigler, and D.M. Strauss. Convergence Properties of Simple Genetic Algorithms. Technical Report 159, Logic of Computers Group, University of Michigan, July 1974.
- [4] Bledsoe, W.W. and I. Browning. Pattern Recognition and Reading by Machine. In Proceedings of the Eastern Joint Computer Conference, 1959.
- [5] Cavicchio, D.J. Adaptive Search Using Simulated Evolution. Technical Report 03296-4-T, Computer and Communication Sciences Department, University of Michigan, August 1970.
- [6] De Jong, K.A. Analysis of the Behavior of a Class of Genetic Adaptive Systems. Technical Report 185, Logic of Computers Group, University of Michigan, August 1975.
- [7] Frantz, D.R. Non-linearities in Genetic Adaptive Search. Technical Report 138, Logic of Computers Group, University of Michigan, September 1972.
- [8] Gaines, B.R. Axioms for Adaptive Behavior. International Journal of Man-Machine Studies 4: 169-199, 1972.
- [9] Holland, J.H. Adaptation in Natural and Artificial Systems. The University of Michigan Press, 1975.
- [10] Hollstien, R.B. Artificial Genetic Adaptation in Computer Control Systems. Technical Report 03296-14-T, Computer and Communication Sciences Department, University of Michigan, April 1971.
- [11] Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. E.A. Feldman and J. Feldman (eds.) Computers and Thought. McGraw-Hill, 1963, pp. 71-105.

- [12] Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. II -- Recent Progress. IBM Journal of Research and Development 11: 601-617, 1967.
- [13] Tsypkin, Y.Z. Adaptation and Learning in Automatic Systems. Academic Press, 1971.
- [14] Uhr, L. and C. Vossler. A Pattern-Recognition Program that Generates, Evaluates, and Adjusts its Own Operators. E.A. Feigenbaum and J. Feldman (eds.) Computers and Thought. McGraw-Hill, 1963, pp. 251-268.